SIAM J. NUMER. ANAL.
Vol. x, No. x, pp. 1–23, x x

CONVERGENCE PAST SINGULARITIES FOR A FULLY DISCRETE APPROXIMATION OF CURVATURE DRIVEN INTERFACES*

RICARDO H. NOCHETTO† AND CLAUDIO VERDI‡

Abstract. Consider a closed surface in $\mathbb{R}^n$ of codimension 1 which propagates in the normal direction with velocity proportional to its mean curvature plus a forcing term. This geometric problem is first approximated by a singularly perturbed parabolic double obstacle problem with small parameter $\varepsilon > 0$. Conforming piecewise linear finite elements over a quasi-uniform and strongly acute mesh of size h are further used for space discretization, and combined with backward differences for time discretization with uniform time-step τ. It is shown that the zero level set of the fully discrete solution converges past singularities to the true interface, provided $\tau, h^2 \approx o(\varepsilon^3)$ and no fattening occurs. If the more stringent relations $\tau, h^2 \approx \mathcal{O}(\varepsilon^4)$ are enforced, then a linear rate of convergence $\mathcal{O}(\varepsilon)$ for interfaces is derived in the vicinity of regular points, namely those for which the underlying viscosity solution is nondegenerate. Singularities and their smearing effect are also studied. The analysis is based on constructing discrete barriers via a parabolic projection, Lipschitz dependence of viscosity solutions with respect to perturbations of data, and discrete nondegeneracy. These issues are proven, along with quasi-optimality in 2D of the parabolic projection in L^∞ with respect to both order and regularity requirements for functions in $W_p^{2,1}$.

Key words. interface evolution, viscosity solutions, reaction-diffusion, double obstacle, singular perturbation, finite elements, discrete maximum principle, pointwise error estimates, discrete nondegeneracy, interface convergence

AMS(MOS) subject classifications. 35D05, 53A10, 35K57, 35B85, 35B25, 65M60, 35B50, 65M12, 65M15

1. Introduction.

Let $\Sigma(t)$ be a closed surface in $\mathbb{R}^n$ of codimension 1 that emanates from a smooth surface Σ_0 and propagates in its normal direction with velocity V (positive if pointing inwards) satisfying the geometric law

$$V = \kappa + g. \tag{1.1}$$

Hereafter κ denotes the sum of the principal curvatures of $\Sigma(t)$ (positive if the enclosed set by $\Sigma(t)$ is locally mean convex), and g is a bounded given driving force. Besides its intrinsic geometric interest, (1.1) may be viewed as a scaled version of the interface law for physical processes varying from crystal growth to flame propagation and two-phase fluid flow. In those cases additional PDEs for the bulk couple with (1.1) and give rise to g. Singularity formation and continuation of the flow past singularities is then a relevant physical issue. Convergence past singularities of fully discrete methods is a companion scientific issue of paramount importance, most notably in regard to the reliable computation of more complex flows. This is still an open problem for all fully discrete approximations of (1.1).

The purely geometric evolution (1.1) already exhibits singularities and topological changes, and even fattening may occur [1], [3], [36]. Therefore the classical differential geometry approach breaks down past singularities. The level set approach overcomes this difficulty and provides a meaningful notion of evolution past singularities by

* This work was partially supported by NSF Grant DMS-9305935, MURST, and CNR Contracts 92.00833.01, 93.00564.01.

† Department of Mathematics and Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742, USA.

‡ Dipartimento di Matematica, Università di Milano, 20133 Milano, Italy.

representing $\Sigma(t)$ as the zero level set of an auxiliary continuous function $w(\cdot,t)$, all whose level sets evolve formally in their normal direction according to (1.1) [31]. Such a w is thus the unique viscosity solution to the degenerate nonlinear PDE

$$\partial_t w - |\nabla w| \operatorname{div}\left(\tfrac{\nabla w}{|\nabla w|}\right) - |\nabla w| g = 0 \quad \text{in } \mathbb{R}^n \times (0,T), \tag{1.2}$$

satisfying the initial condition $w(\cdot,0) = w_0(\cdot) \in C(\mathbb{R}^n)$ [7], [10], [13], [17]. The zero level set of w_0 coincides with Σ_0 and w_0 is constant outside a large ball; thus $w(\mathbf{x},t)$ turns out to be constant for $|\mathbf{x}|$ large [13]. If d_0 denotes the signed distance to Σ_0 which is negative in $I(0)$, the inside of Σ_0, then a typical choice of w_0 is d_0 truncated at ∞. Since the evolution of the zero level set of $w(\cdot,t)$ is independent of the specific form of w_0, for as long as $I(0) = \{\mathbf{x} \in \mathbb{R}^n : w_0(\mathbf{x}) < 0\}$,

$$\Sigma(t) := \{\mathbf{x} \in \mathbb{R}^n : w(\mathbf{x},t) = 0\}, \quad (\Sigma := \{(\mathbf{x},t) \in \mathbb{R}^n \times [0,T] : \mathbf{x} \in \Sigma(t)\}) \tag{1.3}$$

is then called *generalized* geometric evolution and coincides with the classical motion (1.1) before the onset of singularities [2], [7], [13], [17]. The notions of inside $I(t)$ and outside $O(t)$ of $\Sigma(t)$ are thus meaningful in terms of w:

$$I(t) := \{\mathbf{x} \in \mathbb{R}^n : w(\mathbf{x},t) < 0\}, \quad O(t) := \{\mathbf{x} \in \mathbb{R}^n : w(\mathbf{x},t) > 0\}.$$

The flow (1.3) is unique and is well defined even past singularities. These are typically isolated points $(\mathbf{x},t)$ that correspond to $\nabla w(\mathbf{x},t) = 0$ [1], [36], but $\Sigma(t)$ may also develop interior (fattening) when $g \neq 0$ [3]. However for a number of evolutions of interest, e.g. for cylindrically symmetric hypersurfaces, $\Sigma(t)$ is smooth before or between consecutive singularities. With this in mind, we define Σ^* to be the set of *regular* points $(\mathbf{x},t) \in \Sigma$, namely those for which w is C^1 in a space-time neighborhood and $\nabla w(\mathbf{x},t) \neq 0$.

Let Ω be a convex bounded polyhedron of $\mathbb{R}^n$. We will assume that Ω is sufficiently large so as to contain $\Sigma(t)$ for all $t \le T$, for some $T > 0$, and set $Q := \Omega \times (0,T)$. This is guaranteed either for forcing term $g = 0$, provided Ω contains the convex hull of Σ_0 [13], or for T small. Let g satisfy $g \in W^{2,1}_\infty(Q)$.

The main purpose of this paper is to show interface convergence *past singularities* for a practical fully discrete approximation of (1.3). We first introduce a singularly perturbed parabolic problem [6], [25]–[28], [30]

$$\varepsilon \partial_t u_\varepsilon - \varepsilon \Delta u_\varepsilon + \tfrac{1}{\varepsilon}\psi(u_\varepsilon) \ni \tfrac{c_0}{2} g \quad \text{in } Q, \tag{1.4}$$

subject to $u_\varepsilon(\cdot,0) = \phi_\varepsilon(\cdot)$ in Ω and $u_\varepsilon = 1$ on $\partial\Omega \times (0,T)$, with double obstacle potential Ψ with wells of equal depth located at ± 1

$$\Psi(s) := \begin{cases} 1 - s^2 & \text{if } s \in [-1,1] \\ +\infty & \text{if } s \notin [-1,1], \end{cases} \qquad \psi(s) := \tfrac{1}{2}\Psi'(s) = \begin{cases} (-\infty, 1] & \text{if } s = -1 \\ -s & \text{if } |s| < 1 \\ [-1, \infty) & \text{if } s = 1, \end{cases}$$

and scaling factor $c_0 := \int_{-1}^{1} \Psi^{1/2}(s) ds = \pi/2$. The solution u_ε varies between ± 1, the obstacles, within a narrow transition region of thickness $\mathcal{O}(\varepsilon)$, the noncoincidence set, and attains the values ± 1 elsewhere irrespective of g. Hence, in contrast to (1.2), (1.4) retains the local structure of (1.1). This is extremely important numerically, and one of the reasons why we explore this idea as an alternative to the global level set approach (1.2) [31]. It is not obvious, however, that such a property would be preserved by any full discretization. We will show below that this is actually the case.

Conforming piecewise linear finite elements are further used to discretize (1.4) in space, and backward differences to discretize (1.4) in time. The finite element mesh is supposed to be quasi-uniform and *strongly acute*, which in $\mathbb{R}^2$ means that for all pair of triangles sharing a side the sum of the opposite angles to that side does not exceed $\pi - \theta$ for a constant $\theta > 0$. The resulting *implicit* scheme satisfies the *discrete maximum principle* (DMP), our key tool, and its implementation is straightforward. In fact, the underlying PDE being linear, the unilateral constraints are the only source of nonlinearity. Since no numerical integration is considered, this method represents the best scenario in terms of convergence and error analysis, and is to be compared with the *dynamic mesh algorithm* of [25], [29]. The latter resorts to highly graded meshes and associated space-time dependent relaxation parameter $\varepsilon(\mathbf{x}, t)$, uses an explicit adaptive time stepping, and performs mass lumping on the rather singular terms $\varepsilon\partial_t u_\varepsilon$ and $\psi(u_\varepsilon)/\varepsilon$. These features result in enhanced singularity resolution and efficiency. We expect to study this method in light of the best possible results of the present analysis for a simpler, yet practical, fully discrete approximation.

Let $h > 0$ denote the mesh-size and $\tau > 0$ the time-step. Let $U_{\varepsilon,h,\tau}$ indicate the fully discrete solution, which is continuous piecewise linear in space and time, and $\Sigma_{\varepsilon,h,\tau}$ designate its zero level set, namely

$$\Sigma_{\varepsilon,h,\tau}(t) := \{\mathbf{x} \in \Omega : U_{\varepsilon,h,\tau}(\mathbf{x}, t) = 0\}.$$

We demonstrate convergence of $\Sigma_{\varepsilon,h,\tau}(t)$ to $\Sigma(t)$ past singularities provided no fattening occurs. We now discuss briefly the main ingredients of our rather technical construction of fully discrete barriers and their use. From now on we will denote by C a positive constant independent of the relevant parameters ε, h, τ, and p below, that may change at the various occurrences, and use the symbols $\approx, \lessapprox$, and $\gtrapprox$ to indicate the relations $=, \leq$, and $\geq$ up to factors $|\log h|$.

(I) Continuous barriers. Let g in (1.2) and (1.4) be replaced by $g + \alpha$, and w_0 in (1.2) be replaced by $w_0 + C\alpha$, with $0 < \alpha = o(1)$ to be chosen. Let w^+ be the corresponding solution of (1.2) and u^+ that of (1.4). Using w^+, we construct in §3 an *explicit* function v^+ and give a formal proof that v^+ is a viscosity supersolution of (1.4), whence

$$u_\varepsilon \leq u^+ \leq v^+.$$

We refer to [30] for a rigorous proof.

(II) Parabolic projection. Given a function $y \in W^{2,1}_p(Q)$ with $p < \infty$, we indicate with $H_{h,\tau}(y)$ the approximation of y obtained via the discrete heat operator with continuous piecewise linear finite elements in space and backward differences in time. We prove in §4 the 2D pointwise error estimate

$$\|y - H_{h,\tau}(y)\|_{L^\infty(Q)} \lessapprox Cp^2\Big(h^{2-4/p} + \tau^{1-2/p}\Big)\|y\|_{W^{2,1}_p(Q)},$$

which is optimal in both order of convergence (up to logarithmic factors) and regularity requirements, and extends that in [11] valid for $y \in W^{2,1}_\infty(Q)$; note that $y = u^+ \notin W^{2,1}_\infty(Q)$. This result constitutes our sole restriction to 2D, but it is likely to hold in higher dimensions (see note added in proof).

(III) Discrete nondegeneracy. We demonstrate in §6 that $U_{\varepsilon,h,\tau}$ has a prescribed quadratic growth away from the fully discrete free boundary, which for all $(\mathbf{x}, t) \in \partial\{U_{\varepsilon,h,\tau} > -1\}$ means

$$\max_{|\mathbf{y}-\mathbf{x}|^2+(t-s)\leq r^2} U_{\varepsilon,h,\tau}(\mathbf{y}, s) \geq -1 + C\tfrac{r^2}{\varepsilon^2}.$$

We then proceed as follows. We first project u^+ using (II) and next lift $H_{h,\tau}(u^+)$ by the small amount $Ch^2\varepsilon^{-2} \gtrapprox \|u^+ - H_{h,\tau}(u^+)\|_{L^\infty(Q)}$, for the resulting function $V^+_{\varepsilon,h,\tau}$ to be a candidate for discrete supersolution. We must however compensate for the antimonotone nature of the potential $\psi(u_\varepsilon)/\varepsilon = -u_\varepsilon/\varepsilon$ in (1.4), which gives rise to an error $\approx Ch^2\varepsilon^{-3}$; this is an essential difference with respect to [14] that examines a monotone problem. Such a compensation is accomplished by perturbing g by the amount $\alpha \approx Ch^2\varepsilon^{-3}$, which has the implicit effect of a horizontal shift without geometric constraints on the mesh. We conclude that the parabolic relations

$$h \approx o(\varepsilon^{3/2}), \quad \tau \approx o(\varepsilon^3) \tag{1.5}$$

imply that $V^+_{\varepsilon,h,\tau}$ is a discrete supersolution, and thus $V^+_{\varepsilon,h,\tau} \geq U_{\varepsilon,h,\tau}$, for $\alpha = o(1)$. Exploiting the continuous dependence of w with respect to g, this is next used to infer convergence of level sets, a substantially more subtle issue than convergence of solutions [5], [23], [33]. If $\mathbf{x} \in I(t)$, then $w^+(\mathbf{x}, t) < 0$ for α small, and thus $v^+ = -1$ by the explicit construction (I), in a vicinity of $(\mathbf{x}, t)$. This in turn yields

$$U_{\varepsilon,h,\tau} \leq V^+_{\varepsilon,h,\tau} \lessapprox u^+ + C\tfrac{h^2}{\varepsilon^2} \leq v^+ + C\tfrac{h^2}{\varepsilon^2} = -1 + C\tfrac{h^2}{\varepsilon^2}.$$

We finally deduce $U_{\varepsilon,h,\tau}(\mathbf{x}, t) = -1$ for otherwise (III) would lead to a contradiction. This is in essence our convergence result which, although it does not provide a rate, seems to be the first one for discrete interfaces valid after the onset of singularities. Moreover, it is to be compared with Γ-convergence in L^1 for stationary solutions, i.e. solutions to prescribed mean curvature problems, which requires $h = o(\varepsilon)$ but asserts nothing about convergence of interfaces [4]. In this light, relations (1.5) are probably sharp, but such a claim deserves further investigation. We would like to point out the importance of dealing with maximum norm estimates which lead to convergence of interfaces, the only objects of interest for the geometric problem at hand.

Discrete interface error estimates might be expected for regular evolutions. Rather surprisingly, we can also prove a linear rate of convergence $\mathcal{O}(\varepsilon)$ in the vicinity of regular points $(\mathbf{x}, t) \in \Sigma^*$, even past singularities, as a consequence of the following key properties of §2:

(IV) Lipschitz properties: $\|w^+ - w\|_{L^\infty(Q)} \leq C\alpha, \quad \|w^+\|_{W^{1,1}_\infty(Q)} \leq C.$

In fact, if $\nabla w(\mathbf{x}, t) \neq 0$ then $w(\mathbf{x} - \delta\mathbf{n}, t) \leq -C\varepsilon$ with $\mathbf{n} = \nabla w(\mathbf{x}, t)/|\nabla w(\mathbf{x}, t)|$ being a unit normal to $\Sigma(t)$ at $\mathbf{x}$ and $\delta = C\varepsilon/|\nabla w(\mathbf{x}, t)|$, whence (IV) yields

$$w^+(\mathbf{x} - \delta\mathbf{n}, t) \leq w(\mathbf{x} - \delta\mathbf{n}, t) + C\alpha \leq -\varepsilon,$$

provided $\alpha = C\varepsilon$. The Lipschitz regularity in (IV) shows that $w^+ < 0$, and thus $v^+ = -1$ by the explicit construction (I), in an ε-vicinity of $(\mathbf{x} - \delta\mathbf{n}, t)$. Upon arguing as above but with $\alpha = C\varepsilon$, we infer that $U_{\varepsilon,h,\tau}(\mathbf{x} - \delta\mathbf{n}, t) = -1$ and

$$\operatorname{dist}(\mathbf{x}, \Sigma_{\varepsilon,h,\tau}(t)) \leq C\varepsilon|\nabla w(\mathbf{x}, t)|^{-1} \quad \forall\ \varepsilon \leq \varepsilon^*, (\mathbf{x}, t) \in \Sigma^\sharp,$$

with ε^* depending only on the compact subset $\Sigma^\sharp$ of Σ^* and not on the specific $(\mathbf{x}, t)$, under the more severe (parabolic) constraints

$$h \approx \mathcal{O}(\varepsilon^2), \quad \tau \approx \mathcal{O}(\varepsilon^4).$$

An $\mathcal{O}(\varepsilon)$ estimate for the discrete transition layer thickness emerges as a by-product. These ideas are also used in understanding how the discrete problem behaves near certain singularities and demonstrating that the discrete transition layer thickness becomes $\mathcal{O}(\varepsilon^{1/2})$. This quantifies a smearing effect observed in computations [25], [29], [32].

This paper is organized as follows. In §2 we prove Lipschitz continuity of w and show Lipschitz dependence on data of both w and $\Sigma(t)$. In §3 we pose (1.4) as a variational inequality, derive crucial global regularity properties and construct several continuous barriers. In §4 we prove a convenient version of the DMP, along with a quasi-optimal L^∞ error estimate for the parabolic projection with underlying regularity $W_p^{2,1}(Q)$. Both results are instrumental in constructing fully discrete barriers in §5. We demonstrate discrete nondegeneracy in §6 and derive convergence of interfaces in §7. We finally prove error estimates for regular points in §8 and elaborate upon singularities and their smearing effect in §9.

2. Lipschitz dependence on data. It will turn out to be essential to compare *viscosity* solutions w for perturbed data g and w_0. We will need more than the usual continuous dependence. In fact we will use Lipschitz dependence for solutions on the maximum norm which, under a further nondegeneracy condition, yields Lipschitz dependence for interfaces in distance. This is dictated by the following Lipschitz regularity result, which would just be a consequence of translation invariance of (1.2) if g were constant.

THEOREM 2.1. *There exists a constant $C > 0$ depending only on $\|w_0\|_{W^2_\infty(\Omega)}$ and $\|g\|_{W^{1,1}_\infty(Q)}$ such that*

$$\|w\|_{W^{1,1}_\infty(Q)} \le C e^{t\|\nabla g\|_{L^\infty(Q)}}.$$

Proof. The proof is just an application of the maximum principle to suitable auxiliary functions. For $0 < \beta \ll 1$ and $\mathbf{p} = \{p_i\}_{i=1}^n$, set

$$a_{ij}(\mathbf{p}) := (1+\beta)\delta_{ij} - \frac{p_i p_j}{|\mathbf{p}|_\beta^2}, \quad |\mathbf{p}|_\beta := \left(|\mathbf{p}|^2 + \beta^2\right)^{1/2},$$

and note that the symmetric matrix $\{a_{ij}(\mathbf{p})\}_{i,j=1}^n$ is positive definite uniformly in $\mathbf{p}$ because $\sum_{i,j=1}^n a_{ij}(\mathbf{p}) s_i s_j \ge (\beta + \beta^2/|\mathbf{p}|_\beta^2)|\mathbf{s}|^2 > \beta|\mathbf{s}|^2$ for all $\mathbf{s} \in \mathbb{R}^n$.

Since w is constant for $\mathbf{x}$ large [13], g can be assumed to possess compact support in $\mathbb{R}^n$. Let $z = z_\beta \in C^\infty(\mathbb{R}^n \times (0,T))$ be the classical solution to the following (quasilinear) uniformly parabolic regularized version of (1.2) [13]:

$$\partial_t z = \sum_{i,j=1}^n a_{ij}(\nabla z)\ \partial^2_{x_i x_j} z + |\nabla z|_\beta\ g, \tag{2.1}$$

subject to the initial condition $z(\cdot, 0) = w_{0,\beta}(\cdot)$, a regularization of w_0. Classical PDE theory yields existence and uniqueness of a bounded z along with bounded first derivatives in space and time, with bounds that possibly depend on β [21, p.558]. Such a z is also known to converge uniformly on compact sets to w as $\beta \downarrow 0$ [7], [13], [17]. Our intention is to prove Lipschitz regularity of z for $w_{0,\beta} \in W^2_\infty(\Omega)$ uniformly in β, which immediately yields the asserted regularity of w.

We show first that z is Lipschitz continuous in space. For $G := \|\nabla g\|_{L^\infty(Q)}$ consider the auxiliary smooth (bounded) positive function

$$\zeta(\mathbf{x}, t) := \left(|\nabla z(\mathbf{x}, t)|^2 + \beta^2\right) e^{-2Gt} = |\nabla z(\mathbf{x}, t)|_\beta^2 e^{-2Gt}.$$

We claim ζ is bounded only in terms of w_0. Since $\partial_t \zeta = 2e^{-2Gt} \sum_{k=1}^n \partial_{x_k} z\ \partial_t \partial_{x_k} z - 2G\zeta$, we deduce the following linear uniformly parabolic PDE for ζ with smooth bounded coefficients:

$$\partial_t\zeta - \sum_{i,j=1}^{n} a_{ij}(\nabla z)\ \partial^2_{x_ix_j}\zeta + \sum_{m=1}^{n}\Big(\sum_{i,j=1}^{n}\partial_{p_m}a_{ij}(\nabla z)\ \partial^2_{x_ix_j}z + \partial_{p_m}|\nabla z|_\beta\ g\Big)\partial_{x_m}\zeta$$

$$= -2\zeta\big(G + \nabla g\cdot\tfrac{\nabla z}{|\nabla z|_\beta}\big) - 2e^{-2Gt}\sum_{i,j,k=1}^{n} a_{ij}(\nabla z)\ \partial^2_{x_ix_k}z\ \partial^2_{x_jx_k}z,$$

which we write as $\partial_t\zeta - \mathcal{L}\zeta = \mathrm{I} + \mathrm{II}$. We readily have $\mathrm{I} \le 2\zeta(|\nabla g| - G) \le 0$. Let now $\mathbf{R} = \{r_{ij}\}_{i,j=1}^n$ be an orthogonal matrix that diagonalizes the Hessian of z at $(\mathbf{x},t)$. That is, if $z(\mathbf{x},t) = Z(\mathbf{y},t) = Z(R\mathbf{x},t)$, then $\partial^2_{x_ix_k}z = \sum_{m=1}^n r_{im}\ \partial^2_{y_my_m}Z\ r_{km}$. By virtue of the orthogonality of R and the fact that $\{a_{ij}\}_{i,j}^n$ is positive definite, we readily conclude

$$-\tfrac12 e^{2Gt}\mathrm{II} = \sum_{i,j,k=1}^{n} a_{ij}(\nabla z)\Big(\sum_{m=1}^{n} r_{im}\ \partial^2_{y_my_m}Z\ r_{km}\Big)\Big(\sum_{l=1}^{n} r_{jl}\ \partial^2_{y_ly_l}Z\ r_{kl}\Big)$$

$$= \sum_{i,j,m,l=1}^{n} a_{ij}(\nabla z)\ r_{im}\ r_{jl}\ \partial^2_{y_my_m}Z\ \partial^2_{y_ly_l}Z\sum_{k=1}^{n} r_{km}\ r_{kl}$$

$$= \sum_{m=1}^{n}\Big(\sum_{i,j=1}^{n} r_{im}\ a_{ij}(\nabla z)\ r_{jm}\Big)\big(\partial^2_{y_my_m}Z\big)^2 \ge 0.$$

We thus infer that $\partial_t\zeta - \mathcal{L}\zeta \le 0$. Since $\zeta(\cdot,0) \le \|\nabla w_{0,\beta}\|^2_{L^\infty(\Omega)} + \beta^2 =: W$, and ζ is bounded, the classical comparison principle for linear uniformly parabolic PDE [34, p.183] yields $0 < \zeta(\mathbf{x},t) \le W$ or equivalently the asserted estimate

$$|\nabla z(\mathbf{x},t)|^2 \le \big(\|\nabla w_{0,\beta}\|^2_{L^\infty(\Omega)} + \beta^2\big)e^{2Gt} - \beta^2 \quad \forall\ (\mathbf{x},t)\in Q.$$

To prove Lipschitz continuity of z in time, uniformly in β, we differentiate (2.1) with respect to t, and so obtain the linear parabolic PDE for $\varphi(\mathbf{x},t) := \partial_t z(\mathbf{x},t)e^{-Gt}$:

$$\partial_t\varphi - \mathcal{L}\varphi + G\varphi = e^{-Gt}|\nabla z|_\beta\partial_t g = \zeta^{1/2}\partial_t g =: \mathrm{III}.$$

Since $|\mathrm{III}| \le W^{1/2}\|\partial_t g\|_{L^\infty(Q)}$ and $\|\varphi(\cdot,0)\|_{L^\infty(Q)}$ is dictated only by $\|w_{0,\beta}\|_{W^2_\infty(\Omega)}$ and $\|g\|_{W^{1,1}_\infty(Q)}$, the classical comparison principle thus yields the asserted estimate

$$|\partial_t z(\mathbf{x},t)| = |\varphi(\mathbf{x},t)|e^{Gt}$$
$$\le \Big(\|\varphi(\cdot,0)\|_{L^\infty(Q)} + \tfrac{1}{G}\big(\|\nabla w_{0,\beta}\|^2_{L^\infty(\Omega)} + \beta^2\big)^{1/2}\|\partial_t g\|_{L^\infty(Q)}\Big)e^{Gt}. \qquad \square$$

Let now $0 < \alpha = o(1)$, and let $w^\pm$ and $\Sigma^\pm$ be the viscosity solutions and their zero level sets corresponding to data $g \pm \alpha$ and $w_0 \pm C\alpha$.

Corollary 2.1. $\quad w - C\alpha \le w^- \le w \le w^+ \le w + C\alpha \quad$ *in* Q.

Proof. Let $v^+(\cdot,t) := w(\cdot,t) + \alpha\sigma(t)$ with σ to be chosen in such a way that $v = v^+$ will be a supersolution of (1.2) with perturbed data $g+\alpha$ and $w_0 + C\alpha$. Since σ does not depend on $\mathbf{x}$, we can write in the viscosity sense

$$\partial_t v = |\nabla v|\mathrm{div}\big(\tfrac{\nabla v}{|\nabla v|}\big) + |\nabla v|g + \alpha\sigma'(t) \ge |\nabla v|\mathrm{div}\big(\tfrac{\nabla v}{|\nabla v|}\big) + |\nabla v|(g+\alpha),$$

provided $\sigma'(t) = \|\nabla v\|_{L^\infty(Q)} = \|\nabla w\|_{L^\infty(Q)}$, which is feasible by virtue of Theorem 2.1. If also $\sigma(0) = C$, then $v(\cdot,0) = w_0(\cdot) + \alpha\sigma(0) \ge w^+(\cdot,0)$, whence $w \le w^+ \le v$ in Q by a comparison result for supersolutions of (1.2) [7], [13], [17]. Since $0 \le \sigma \le C$, we conclude $w \le w^+ \le v \le w + C\alpha$. $\square$

We recall that Σ^* is the set of *regular* points $(\mathbf{x},t)\in\Sigma$, namely those for which w is C^1 in a space-time neighborhood and $\nabla w(\mathbf{x},t)\neq 0$. Since Σ_0 is smooth then Σ^* extends all the way to $t=0$.

COROLLARY 2.2. *Let $\Sigma^\sharp$ be a compact subset of Σ^*. There exist constants $\alpha^*>0$ depending on $\Sigma^\sharp$ and $C>0$ depending on data such that for all $\alpha\le\alpha^*$*

$$\mathrm{dist}(\mathbf{x},\Sigma^\pm(t))\le\tfrac{C}{|\nabla w(\mathbf{x},t)|}\alpha\quad\forall\ (\mathbf{x},t)\in\Sigma^\sharp.$$

Proof. Since w is C^1 in a neighborhood of $(\mathbf{x},t)$ we can write

$$w(\mathbf{x}-\delta\mathbf{n},t)=w(\mathbf{x},t)-\delta\nabla w(\mathbf{x},t)\cdot\mathbf{n}+o(\delta)\le-\tfrac12\delta|\nabla w(\mathbf{x},t)|,$$

for $\mathbf{n}:=\nabla w(\mathbf{x},t)/|\nabla w(\mathbf{x},t)|$ and $0<\delta\le D$ with $D\ll 1$ depending on the modulus of continuity of ∇w and a lower bound $W>0$ for $|\nabla w|$. Since $\Sigma^\sharp$ is compact, these quantities depend on $\Sigma^\sharp$ but not on $(\mathbf{x},t)$. With the aid of Corollary 2.1 we infer that

$$w^+(\mathbf{x}-\delta\mathbf{n},t)\le w(\mathbf{x}-\delta\mathbf{n},t)+C\alpha\le-\tfrac12\delta|\nabla w(\mathbf{x},t)|+C\alpha=0,$$

provided $\delta=2C\alpha/|\nabla w(\mathbf{x},t)|\le D$, which is guaranteed for $\alpha\le\alpha^*:=DW/2C$. Since also $w^+(\mathbf{x},t)\ge w(\mathbf{x},t)=0$, and w^+ is continuous, we realize that $w^+(\cdot,t)$ has a zero on the straight segment joining $\mathbf{x}-\delta\mathbf{n}$ and $\mathbf{x}$, thereby completing the proof. $\square$

The idea behind Corollary 2.2 goes back to [5], and expresses the geometric fact that nondegeneracy together with maximum norm error estimates for solutions lead to interface error control. We will see later that, because of the lack of meaningful error estimates for solutions of (1.4) for large time, this idea does not apply to (1.4) nor to its fully discrete counterpart.

3. Parabolic variational inequality. Let $\varepsilon>0$ be the relaxation parameter, $\mathcal{K}$ be the convex set

$$\mathcal{K}:=\{\varphi\in H^1(\Omega):|\varphi|\le 1,\varphi=1\text{ on }\partial\Omega\},$$

and $\phi_\varepsilon\in W^2_\infty(\Omega)\cap\mathcal{K}$ satisfy $\|\phi_\varepsilon\|_{W^2_\infty(\Omega)}\le C\varepsilon^{-2}$. The double obstacle problem (1.4) can be written equivalently as follows: $u_\varepsilon\in L^\infty(0,T;\mathcal{K})\cap H^1(0,T;L^2(\Omega))$, $u_\varepsilon(\cdot,0)=\phi_\varepsilon(\cdot)$ in Ω, and, for all $\varphi\in\mathcal{K}$ and a.e. $t\in(0,T)$,

$$\varepsilon\langle\partial_t u_\varepsilon,\varphi-u_\varepsilon\rangle+\varepsilon\langle\nabla u_\varepsilon,\nabla(\varphi-u_\varepsilon)\rangle-\tfrac{1}{\varepsilon}\langle u_\varepsilon,\varphi-u_\varepsilon\rangle\ge\tfrac{c_0}{2}\langle g,\varphi-u_\varepsilon\rangle. \tag{3.1}$$

Hereafter we will use the notation $\langle\varphi_1,\varphi_2\rangle:=\int_\Omega\varphi_1\varphi_2$ for any $\varphi_1,\varphi_2\in L^2(\Omega)$. This problem admits a unique solution, which satisfies the following *optimal* a priori bound. This bound is also valid for any solution of (3.1) with forcing of order $\mathcal{O}(\varepsilon^{-1})$.

LEMMA 3.1. $|u_\varepsilon|_{W^{2,1}_p(Q)}=\|D^2_{\mathbf{x}}u_\varepsilon\|_{L^p(Q)}+\|\partial_t u_\varepsilon\|_{L^p(Q)}\le C\ p\ \varepsilon^{-2}$ *as $p\uparrow\infty$.*

Proof. It suffices to rewrite (3.1) as

$$\langle\partial_t u_\varepsilon,\varphi-u_\varepsilon\rangle+\langle\nabla u_\varepsilon,\nabla(\varphi-u_\varepsilon)\rangle\ge\langle\tfrac{1}{\varepsilon^2}u_\varepsilon+\tfrac{c_0}{2\varepsilon}g,\varphi-u_\varepsilon\rangle,$$

and use $W^{2,1}_p$ regularity theory for the usual parabolic obstacle problem associated with the heat operator with right-hand side $f_\varepsilon:=(u_\varepsilon+\frac{c_0}{2}\varepsilon g)\varepsilon^{-2}=\mathcal{O}(\varepsilon^{-2})$ and initial datum ϕ_ε, which is a consequence of linear theory [15, p.77].

The dependence on p comes also from linear theory and can be determined as follows. Let u be the solution to $\partial_t u-\Delta u=f\in L^\infty(Q)$ with vanishing Dirichlet condition and $u(\cdot,0)=\phi\in W^2_\infty(\Omega)$, where ϕ satisfies suitable compatibility conditions at $\partial\Omega$. Let $\sigma(t)$ be a smooth function such that $\sigma(0)=1$ and vanishes for large t. The auxiliary function $v(\mathbf{x},t):=\phi(\mathbf{x})\sigma(t)$ satisfies $\|v\|_{W^{2,1}_\infty(Q)}\le C$, so we can replace u by $u-v$ in the asserted estimate and therefore deal with zero initial trace. Suppose now that $\Omega = \mathbb{R}^n$. The solution u is just the convolution of the heat kernel and f. Calderón-Zygmund theory of parabolic singular integrals then yields the desired linear dependence on p [21]. Otherwise, if Ω is bounded but smooth, then we localize near $\partial\Omega$ and flatten $\partial\Omega$. The problem thus becomes an assertion about the behavior in the half space of parabolic problems with smooth coefficients and zero Dirichlet condition. Extending f oddly to the entire space $\mathbb{R}^n \times (0,T)$, the odd extension of u solves a parabolic problem in $\mathbb{R}^n \times (0,T)$ for which the previous reasoning applies. The argument as usual concludes by pasting together the various contributions due to localization, which happens to be a procedure independent of p. $\square$

We stress that u_ε is not in $W^{2,1}_\infty(Q)$ with norm of order $\mathcal{O}(\varepsilon^{-2})$ because this would require higher regularity of $f_\varepsilon = (u_\varepsilon + \frac{c_0}{2}\varepsilon g)\varepsilon^{-2} = \mathcal{O}(\varepsilon^{-2})$, and so of u_ε still with norms of order $\mathcal{O}(1)$, which we cannot expect. The following simple result quantifies how far $\partial_t u_\varepsilon$ and $D^2_{\mathbf{x}} u_\varepsilon$ are from being bounded.

LEMMA 3.2. *Let $y \in W^{2,1}_p(Q)$ be so that $\|y\|_{W^{2,1}_p(Q)} \le Bp$ for all $2+n < p < \infty$, and let $\omega(s) := CBs|\log s|$ for $s > 0$ and $C > 0$ sufficiently large. Then we have*

$$\|\nabla y(\mathbf{x}_1,\cdot) - \nabla y(\mathbf{x}_2,\cdot)\|_{L^\infty(0,T)} \le \omega(|\mathbf{x}_1 - \mathbf{x}_2|) \quad \forall\ \mathbf{x}_1,\mathbf{x}_2 \in \Omega,\ |\mathbf{x}_1 - \mathbf{x}_2| \ll 1,$$
$$\|y(\cdot,t_1) - y(\cdot,t_2)\|_{L^\infty(\Omega)} \le \omega(|t_1 - t_2|) \quad \forall\ t_1,t_2 \in [0,T],\ |t_1 - t_2| \ll 1.$$

Proof. For $y \in W^{2,1}_p(Q)$ with $m := 2+n < p < \infty$, there exists $C > 0$ such that

$$\|y\|_{C^{2-m/p,1-m/2p}(\bar{Q})} \le C\|y\|_{W^{2,1}_p(Q)}, \tag{3.2}$$

with $C > 0$ bounded uniformly in p as $p \uparrow \infty$ [21, pp.8,80]. It follows from (3.2) that

$$\|\nabla y(\cdot,t)\|_{C^{1-m/p}(\bar{\Omega})} \le C\|y\|_{W^{2,1}_p(Q)} \le CBp \quad \forall\ m < p < \infty.$$

Let $\mathbf{x}_1,\mathbf{x}_2 \in \Omega$ satisfy $0 < |\mathbf{x}_1 - \mathbf{x}_2| < e^{-m}$. Then for $p = -\log|\mathbf{x}_1 - \mathbf{x}_2| > m$ we get

$$|\nabla y(\mathbf{x}_1,t) - \nabla y(\mathbf{x}_2,t)| \le CBp|\mathbf{x}_1 - \mathbf{x}_2|^{1-m/p} \le \omega(|\mathbf{x}_1 - \mathbf{x}_2|) \quad \forall\ t \in [0,T],$$

because $s^{-m/|\log s|} = e^m$. Similar reasoning applies for the remaining estimate. $\square$

We next introduce continuous barriers for u_ε. For monotone parabolic variational inequalities, as the one-phase Stefan problem [15], any vertical shift of the actual solution leads to a barrier. Here, because of the antimonotone source ψ, a constant vertical shift of u_ε would not have the same effect. If we allow time dependent shifts, then it is easy to check that an obvious candidate for supersolution would be $u^+_\varepsilon(\cdot,t) := u_\varepsilon(\cdot,t) + \delta e^{t/\varepsilon^2}$ for any $\delta > 0$. Because of the singular nature of (3.1) for ε small, this barrier is meaningless for short time of order $t = \mathcal{O}(\varepsilon^2|\log\delta|)$ in that the discrepancy between u_ε and u^+_ε is $\mathcal{O}(1)$. This behavior is consistent with the short time evolution of (1.4), studied in [6], and makes L^∞ error estimates for fully discrete solutions meaningless. The idea of [5] thus fails.

Since our primary interest lies on the interface $\Sigma_\varepsilon(t) := \{\mathbf{x} \in \Omega : u_\varepsilon(\mathbf{x},t) = 0\}$, rather than $u_\varepsilon(t)$, we replace the vertical shift with a horizontal one centered on the continuous interface $\Sigma(t)$, as in [2], [6], [12], [26], [27], [30]. For the fully discrete problem, this crucial idea cannot be achieved by a simple translation, as in the above papers, but implicitly via a perturbation of the forcing term g. Whether this preserves interface distance control is not obvious, and in fact depends on Corollary 2.2 and the construction of other more explicit barriers $u^\pm$ and $v^\pm$. This topic is discussed next.

Let hereafter $\alpha \ge \varepsilon$. In the sequel w and $w^\pm$ will stand for viscosity solutions to

(1.2) with forcing terms g and $g \pm \alpha$, and initial traces

$$w(\cdot, 0) = d_0(\cdot), \quad w^{\pm}(\cdot, 0) = d_0(\cdot) \pm l,$$

where d_0 is the signed distance to Σ_0 negative in I_0 and, for Λ sufficiently large,

$$l := (\Lambda e^{2(G+1)T} + \pi)\alpha, \quad G := \|\nabla g\|_{L^\infty(Q)}.$$

Let $d^{\pm}(\cdot, t)$ be the signed distance function to $\Sigma^{\pm}(t) = \{\mathbf{x} \in \Omega : w^{\pm}(\mathbf{x}, t) = 0\}$, that is positive in $O^{\pm}(t) := \{\mathbf{x} \in \Omega : w^{\pm}(\mathbf{x}, t) > 0\}$. It is known that d^+ satisfies the following inequality in the viscosity sense, provided $\Sigma^+(t)$ has empty interior [2], [12]:

$$\partial_t d^+(\mathbf{x}, t) - \Delta d^+(\mathbf{x}, t) - g\big(\mathbf{x} - d^+(\mathbf{x}, t)\nabla d^+(\mathbf{x}, t), t\big) - \alpha \geq 0 \quad \text{in } \{d^+ > 0\}. \tag{3.3}$$

In case $\Sigma^+(t)$ fattens up, then $d^+(\mathbf{x}, t)$ must be replaced by $\text{dist}(\mathbf{x}, \partial\{w^+(\cdot, t) < 0\})$ for (3.3) to hold [2, p.446]; (3.3) is also valid with ≤ 0 and $g - \alpha$ for $-\text{dist}(x, \partial\{w^-(\cdot, t) > 0\})$. In the sequel we will use the notation $d^{\pm}(\mathbf{x}, t)$ instead of $\pm\text{dist}(\mathbf{x}, \partial\{w^{\pm}(\cdot, t) \lesseqgtr 0\})$. Let $\iota^{\pm}(\mathbf{x}, t)$ indicate the following stretched variables with time dependent shift $\sigma(t)$

$$\iota^{\pm}(\mathbf{x}, t) := \tfrac{d^{\pm}(\mathbf{x}, t)}{\varepsilon} \mp \sigma(t), \quad \sigma(t) := \tfrac{\pi}{2} + \Lambda e^{2(G+1)(T-t)}.$$

Let $\gamma, \eta \in C^{1,1}(\mathbb{R})$ be given by

$$\gamma(x) := \begin{cases} -1 & \text{if } x < -\frac{\pi}{2} \\ \sin x & \text{if } x \in [-\frac{\pi}{2}, \frac{\pi}{2}] \\ +1 & \text{if } x > \frac{\pi}{2}, \end{cases} \quad \eta(x) := \begin{cases} \frac{1}{2}\big(x\gamma(x) - c_0 + \gamma'(x)\big) & \text{if } |x| \leq \frac{\pi}{2} \\ 0 & \text{if } |x| > \frac{\pi}{2}. \end{cases}$$

They are dictated by formal asymptotics, satisfy $|\eta(x)|, |\eta'(x)| < \gamma'(x)$ for $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$ and solve the ODEs [26], [27], [32]:

$$\gamma''(x) + \gamma(x) = 0, \quad \eta''(x) + \eta(x) = \gamma'(x) - \tfrac{c_0}{2}, \quad \forall\, x \in (-\tfrac{\pi}{2}, \tfrac{\pi}{2}). \tag{3.4}$$

The following function $\Gamma^{\pm}(\cdot; \mathbf{x}, t) \in C^{1,1}(\mathbb{R})$ is an explicit approximation of the traveling wave corresponding to (3.1) with forcing term $g \pm \alpha$ [30]:

$$\Gamma^{\pm}(x; \mathbf{x}, t) := \gamma(x) + \varepsilon\big(g(\mathbf{x}, t) \pm \alpha\big)\eta(x) \quad \forall\, x \in \mathbb{R}, \ (\mathbf{x}, t) \in Q.$$

Note that, for α small, $|\Gamma^{\pm}| \leq 1$, $\Gamma^{\pm}(\cdot; \mathbf{x}, t)$ is strictly increasing in $(-\frac{\pi}{2}, \frac{\pi}{2})$ because $\partial_x \Gamma^{\pm}(x; \cdot, \cdot) > \gamma'(x)/2$, and $\Gamma^{\pm}(x \pm \frac{\pi}{2}; \cdot, \cdot) \gtrless \gamma(x)$. With the aid of (3.4) we also get

$$\partial_x^2 \Gamma^{\pm} + \Gamma^{\pm} - \varepsilon(g \pm \alpha)\big(\partial_x \Gamma^{\pm} - \tfrac{c_0}{2}\big) = \partial_x \Gamma^{\pm} \mathcal{O}(\varepsilon^2) \quad \text{in } (-\tfrac{\pi}{2}, \tfrac{\pi}{2}). \tag{3.5}$$

We recall a result of [30], which has been crucial in showing convergence of $\Sigma_\varepsilon(t)$ to $\Sigma(t)$ as $\varepsilon \downarrow 0$, and give a formal proof to illustrate the above definitions.

LEMMA 3.3. *Let $u^{\pm}$ be the solutions to* (3.1) *with forcing term $g \pm \alpha$ and initial trace $\phi_\varepsilon := \gamma(d_0/\varepsilon)$. Let $v^{\pm}$ be defined by*

$$v^{\pm}(\mathbf{x}, t) := \Gamma^{\pm}(\iota^{\pm}(\mathbf{x}, t); \mathbf{x}, t) = \gamma(\iota^{\pm}(\mathbf{x}, t)) + \varepsilon(g(\mathbf{x}, t) \pm \alpha)\eta(\iota^{\pm}(\mathbf{x}, t)).$$

Then, for Λ sufficiently large, v^+ and v^- are viscosity super and subsolutions of (3.1) *with forcing term $g + \alpha$ and $g - \alpha$ respectively, and they satisfy*

$$v^-(\mathbf{x}, t) \leq u^-(\mathbf{x}, t) \leq u_\varepsilon(\mathbf{x}, t) \leq u^+(\mathbf{x}, t) \leq v^+(\mathbf{x}, t) \quad \forall\, (\mathbf{x}, t) \in Q.$$

Proof. We first set $u = u^+, v = v^+, d = d^+, \iota = \iota^+, \Gamma = \Gamma^+$, and $\Gamma^{(k)} = \partial_x^k \Gamma$.

A standard comparison argument for variational inequalities implies $u \geq u_\varepsilon$, because they satisfy the same problem with right hand sides $g + \alpha \geq g$. To prove this, select $\varphi = \min(u_\varepsilon, u)$ in (3.1) for u_ε, $\varphi = \max(u, u_\varepsilon)$ in (3.1) for u, take the difference of the two equations, and apply Gronwall's lemma.

In order to show that $u \le v$, we have to prove that $\mathcal{J} v := \varepsilon\partial_t v - \varepsilon\Delta v - \frac{1}{\varepsilon} v - \frac{c_0}{2}(g+\alpha) \ge 0$ in $\{v(\mathbf{x},t) < 1\}$, that is $\{\iota(\mathbf{x},t) < \frac{\pi}{2}\}$. Since

$$\begin{aligned} \varepsilon\partial_t v &= \Gamma' \partial_t d - \varepsilon\sigma'(t)\Gamma' + \varepsilon^2 \eta \partial_t g, \\ \varepsilon\Delta v &= \tfrac{1}{\varepsilon}\Gamma'' + \Gamma'\Delta d + 2\varepsilon\eta'\nabla g \cdot \nabla d + \varepsilon^2\eta\Delta g, \end{aligned}$$

we have

$$\begin{aligned} \mathcal{J} v = &- \tfrac{1}{\varepsilon}\big(\Gamma'' + \Gamma - \varepsilon(g+\alpha)(\Gamma' - \tfrac{c_0}{2})\big) + \Gamma'\big(\partial_t d - \Delta d - (g+\alpha)\big) \\ &- \varepsilon\sigma'(t)\Gamma' - 2\varepsilon\eta'\nabla g\cdot\nabla d + \varepsilon^2\eta(\partial_t g - \Delta g). \end{aligned}$$

If $\iota < -\frac{\pi}{2}$, then $\gamma = -1$ and $\eta = 0$, whence $\mathcal{J} v = \frac{1}{\varepsilon} - \frac{c_0}{2}(g+\alpha) > 0$ for ε sufficiently small. If $-\frac{\pi}{2} < \iota < \frac{\pi}{2}$, instead, then $0 < d(\mathbf{x},t) < \varepsilon\big(\sigma(t) + \frac{\pi}{2}\big)$ and

$$\big|g(\mathbf{x},t) - g\big(\mathbf{x} - d(\mathbf{x},t)\nabla d(\mathbf{x},t), t\big)\big| \le \|\nabla g\|_{L^\infty(Q)} d(\mathbf{x},t) \le \varepsilon G\big(\sigma(t) + \tfrac{\pi}{2}\big),$$

whence, in view of (3.3) and property $\Gamma' \ge 0$,

$$\Gamma'\big(\partial_t d - \Delta d - (g+\alpha)\big) \ge -\varepsilon G \Gamma'\big(\sigma(t) + \tfrac{\pi}{2}\big).$$

Using now (3.5), the definition of $\sigma(t)$, and $g \in W^{2,1}_\infty(Q)$ we thus have

$$\mathcal{J} u \ge \varepsilon\Gamma'\big(-G\sigma(t) - \sigma'(t) + \mathcal{O}(1)\big) = \varepsilon\Gamma'\big(G\sigma(t) + \mathcal{O}(1)\big) \ge \varepsilon\Gamma'\big(G\Lambda + \mathcal{O}(1)\big) \ge 0,$$

for a suitable choice of $\Lambda > 0$. Since $\Gamma(\cdot;\mathbf{x},t) \in C^{1,1}(\mathbb{R})$ by construction, we conclude that $\Gamma''\big(\iota(\mathbf{x},t);\mathbf{x},t\big)$ is a bounded function and so a distribution without mass concentrated on $\{\iota(\mathbf{x},t) = -\frac{\pi}{2}\}$. This shows that v is a (formal) viscosity supersolution. It only remains to check the values of v on the parabolic boundary of Q. Since the transition region $\{|v(\cdot,t)| < 1\}$ does not touch $\partial\Omega$, we have $v = 1$ on $\partial\Omega \times (0,T)$. In addition, the facts $d(\mathbf{x},0) = d_0(\mathbf{x}) + l = d_0(\mathbf{x}) + (\frac{\pi}{2} + \sigma(0))\alpha$ and $\alpha \ge \varepsilon$ yield

$$v(\mathbf{x},0) = \Gamma\big(\tfrac{d(\mathbf{x},0)}{\varepsilon} - \sigma(0)\big) \ge \Gamma\big(\tfrac{d_0(\mathbf{x})}{\varepsilon} + \tfrac{\pi}{2}\big) \ge \gamma\big(\tfrac{d_0(\mathbf{x})}{\varepsilon}\big) = \phi_\varepsilon(\mathbf{x}) = u(\mathbf{x},0) \quad \forall\ \mathbf{x} \in \Omega.$$

Therefore we conclude that $v \ge u$ as a consequence of a viscosity comparison result for variational inequalities [30]. The proof is thus complete. ∎

4. Discretization, DMP and parabolic projection. Let $n = 2$ and $\mathcal{D}_h$ denote a quasi-uniform decomposition of $\Omega \subset \mathbb{R}^2$ into triangles of size comparable with $h > 0$. In addition to the usual regularity of $\mathcal{D}_h$ [8], we assume that $\mathcal{D}_h$ is *strongly acute* [9]:

(4.1) *the sum of the opposite angles to the common side of any two adjacent triangles is* $\le \pi - \theta$, *with* $\theta > 0$ *independent of* h.

This is a rather restrictive condition, but convenient for theoretical purposes because it yields a *discrete maximum principle* (DMP) without need of mass lumping. Such a numerical quadrature would otherwise be acting on a singular term, and whether it preserves the desired accuracy is unclear at the moment. This issue is to be investigated further and compared with the present analysis which thus corresponds to the best possible scenario.

Let $\mathcal{V}_h$ be the finite element space of continuous piecewise linear functions over $\mathcal{D}_h$, $\mathcal{V}_h(a) := \{\chi \in \mathcal{V}_h : \chi = a \text{ on } \partial\Omega\}$, and set $\mathcal{K}_h := \mathcal{K} \cap \mathcal{V}_h(1)$. Let Π_h be the usual

Lagrange interpolation operator. Let $\tau := T/N > 0$ be the uniform time-step. Set $t_n := n\tau$ and, for any $\{\chi^n\}_{n=0}^N \subset \mathcal{V}_h$ and χ continuous, respectively,

$$\partial\chi^n := \tfrac{1}{\tau}(\chi^n - \chi^{n-1}) \quad \text{and} \quad \chi^n := \chi(t_n), \quad \bar{\chi}^n := \tfrac{1}{\tau}\int_{t_{n-1}}^{t_n} \chi(\cdot,t)dt \quad (\bar{\chi}^0 := \chi^0).$$

The fully discrete approximation of (3.1) reads: *for* $U^0 := \Pi_h[\gamma(d_0/\varepsilon)] \in \mathcal{K}_h$ *seek* $U^n \in \mathcal{K}_h$, $1 \le n \le N$, *such that, for all* $\chi \in \mathcal{K}_h$,

$$\varepsilon\langle\partial U^n, \chi - U^n\rangle + \varepsilon\langle\nabla U^n, \nabla(\chi - U^n)\rangle - \tfrac{1}{\varepsilon}\langle U^n, \chi - U^n\rangle \ge \tfrac{c_0}{2}\langle g^n, \chi - U^n\rangle. \tag{4.2}$$

We then see that this is a discrete parabolic variational inequality corresponding to the linear operator $\varepsilon\partial_t v - \varepsilon\Delta v - \frac{1}{\varepsilon}v$, and so it is easy to implement. Solvability, and also uniqueness, is guaranteed provided $\tau \le \varepsilon^2$.

Let $\{\phi_i\}_{i=1}^I$ be the canonical basis of $\mathcal{V}_h(0)$. Let $K := \{k_{ij}\}_{i,j=1}^I$ be the stiffness matrix and $M := \{m_{ij}\}_{i,j=1}^I$ be the mass matrix, that is

$$k_{ij} := \langle\nabla\phi_i, \nabla\phi_j\rangle, \quad m_{ij} := \langle\phi_i, \phi_j\rangle.$$

Note that, in view of (4.1), there exists a constant $C_\theta > 0$ depending only on θ such that $k_{ij} \le -C_\theta < 0$ uniformly in h for any pair of adjacent nodes. If $a_{ij} := k_{ij} + \delta\, m_{ij}$, then $a_{ij} \le 0$ provided $i \ne j$ and $0 \le \delta \le C_\theta h^{-2}$; $\{a_{ij}\}_{i,j=1}^I$ is thus an M-matrix.

We will need the following form of DMP. For any $\chi \in \mathcal{V}_h(0)$ set $[\chi]^+ := \Pi_h[\max(\chi, 0)] \in \mathcal{V}_h(0)$ and $[\chi]^- := \Pi_h[\min(\chi,0)] \in \mathcal{V}_h(0)$.

LEMMA 4.1. *For all* $\delta \le C_\theta h^{-2}$ *there holds*

$$\delta\|[\chi]^+\|_{L^2(\Omega)}^2 + \|\nabla[\chi]^+\|_{L^2(\Omega)}^2 \le \delta\langle[\chi]^+, \chi\rangle + \langle\nabla[\chi]^+, \nabla\chi\rangle.$$

Proof. Since $a_{ij} \le 0$ provided $i \ne j$ and $\delta \le C_\theta h^{-2}$, it follows that

$$\begin{aligned}\delta\langle[\chi]^+, \chi - [\chi]^+\rangle + \langle\nabla[\chi]^+, \nabla(\chi - [\chi]^+)\rangle &= \delta\langle[\chi]^+, [\chi]^-\rangle + \langle\nabla[\chi]^+, \nabla[\chi]^-\rangle \\ &= \sum_{i\ne j} a_{ij}[\chi]^+(\mathbf{x}_i)[\chi]^-(\mathbf{x}_j) \ge 0. \qquad \square\end{aligned}$$

A simple application of this lemma is the following parabolic DMP for *noncylindrical* discrete domains D [33]. For each basis function ϕ_i and $1 \le n \le N$, set

$$K_i^n := \mathrm{supp}\phi_i \times [t_{n-1}, t_n].$$

Let D be a (closed) connected union of elements K_i^n, or alternatively the sets of its nodes $(\mathbf{x}_i, t_n)$, and denote by $\partial_p D$ the parabolic boundary of D defined by

$$\partial_p D := \{(\mathbf{x}_i, t_n) \in \partial D : K_i^n \not\subset D \text{ or } (\mathbf{x}_i, t_n) \in \partial_p Q\}.$$

LEMMA 4.2. *Let* $\{\chi^n\}_{n=0}^N \subset \mathcal{V}_h$ *be so that*

$$\langle\partial\chi^n, \phi_i\rangle + \langle\nabla\chi^n, \nabla\phi_i\rangle \le 0 \quad \forall\ (\mathbf{x}_i, t_n) \in D\backslash\partial_p D. \tag{4.3}$$

If $\tau \ge C_\theta^{-1}h^2$, *then*

$$\max_{(\mathbf{x}_i,t_n)\in D} \chi^n(\mathbf{x}_i) \le \max_{(\mathbf{x}_i,t_n)\in\partial_p D} \chi^n(\mathbf{x}_i).$$

Proof. We first point out that (4.3) just relates values $\chi^n(\mathbf{x}_j)$ and $\chi^{n-1}(\mathbf{x}_k)$ for $\mathbf{x}_j, \mathbf{x}_k \in \mathrm{supp}\phi_i$, whence changes of $\{\chi^n\}_{n=0}^N$ outside D will not alter (4.3). On subtracting $\mu := \max\{\chi^n(\mathbf{x}_i) : (\mathbf{x}_i, t_n) \in \partial_p D\}$ from $\{\chi^n\}_{n=0}^N$ we can thus assume

$\mu = 0$ and set $\chi^n(\mathbf{x}_i) = 0$ for all $(\mathbf{x}_i, t_n) \notin D$. Assume by induction that $\chi^{n-1} \le 0$ in Ω, multiply (4.3) by $\max(\chi^n(\mathbf{x}_i), 0)$ and add over $1 \le i \le I$ to get

$$\tfrac{1}{\tau}\langle \chi^n, [\chi^n]^+\rangle + \langle \nabla\chi^n, \nabla[\chi^n]^+\rangle \le \tfrac{1}{\tau}\langle \chi^{n-1}, [\chi^n]^+\rangle \le 0.$$

The desired inequality $\chi^n \le 0$ in Ω, which completes the induction argument, then follows from Lemma 4.1 provided $\tau \ge C_\theta^{-1} h^2$. □

We will also need to approximate functions by piecewise linears preserving differential properties. The following *parabolic projection* will be instrumental in doing so. Given $y \in W_p^{2,1}(Q)$ for $p > 2$ with $y = 0$ on $\partial\Omega \times (0,T)$, let $\{Y^n\}_{n=1}^N \subset \mathcal{V}_h(0)$ be defined by

$$\langle \partial Y^n, \chi\rangle + \langle \nabla Y^n, \nabla\chi\rangle = \langle \partial y^n, \chi\rangle + \langle \nabla \bar{y}^n, \nabla\chi\rangle \quad \forall \chi \in \mathcal{V}_h(0), \tag{4.4}$$

with $Y^0 := \Pi_h(y^0)$. Note that Y^0 makes sense because $y^0 = y(\cdot, 0) \in W_p^{2-2/p}(\Omega) \subset C^0(\bar\Omega)$. Let $\hat{H}_{h,\tau}$ be the parabolic projection operator defined by

$$\hat{H}_{h,\tau}(y;\cdot,t) := \tfrac{t-t_{n-1}}{\tau} Y^n(\cdot) + \tfrac{t_n - t}{\tau} Y^{n-1}(\cdot) \quad \forall\, t_{n-1} < t \le t_n,\ 1 \le n \le N.$$

Continuity in time is immaterial except for the discussion in §9. The following error estimate is a convenient modification of the quasi-optimal L^∞ error estimate shown in [11] in 2D. It is worth stressing that this constitutes the only restriction of our analysis to 2D. Even though such an estimate seems plausible in 3D, its extension is not immediate and thus deserves further thought.

THEOREM 4.1. *If $\tau \ge C^*|\log h|^3 h^2$ for $C^* > 0$ sufficiently large, then*

$$\|y - \hat{H}_{h,\tau}(y)\|_{L^\infty(Q)} \le C p^2 |\log h|^2 \Big(h^{2-4/p} + \tau^{1-2/p}\Big) \|y\|_{W_p^{2,1}(Q)} \quad \textit{as } p \uparrow \infty.$$

In view of (3.2), this error estimate appears to be quasi-optimal in both order and regularity requirement, and thus of some intrinsic interest as well.

The proof of Theorem 4.1 will be a consequence of a series of lemmas. We start by introducing Nitsche's weight function for a given but arbitrary $\mathbf{x}_0 \in \Omega$ [22]:

$$\sigma(\mathbf{x}) := \big(|\mathbf{x} - \mathbf{x}_0|^2 + \omega^2\big)^{1/2} \quad \forall\, \mathbf{x} \in \Omega, \tag{4.5}$$

where $\omega = \omega(h) := Ch|\log h|$ for $C \ge 1$. The following estimate is easy to derive by direct computation.

LEMMA 4.3. $\displaystyle\int_\Omega \sigma(\mathbf{x})^{-s} d\mathbf{x} \le C\big((s-2)\omega^{s-2}\big)^{-1}$ *as $s \downarrow 2$.*

The following interpolation estimate is valid for $r > 2$ and $1 \le s \le 2$ [8, p.122]:

$$\|\varphi - \Pi_h\varphi\|_{L^r(\Omega)} + h\|\nabla(\varphi - \Pi_h\varphi)\|_{L^r(\Omega)} \le Ch^s|\varphi|_{W_r^s(\Omega)}. \tag{4.6}$$

Let P_h and R_h be the L^2 and $H_0^1(\Omega)$ projections onto $\mathcal{V}_h(0)$, namely

$$\begin{aligned} &\varphi \in L^2(\Omega), P_h\varphi \in \mathcal{V}_h(0): \quad \langle P_h\varphi - \varphi, \chi\rangle = 0 \quad \forall\, \chi \in \mathcal{V}_h(0),\\ &\varphi \in H_0^1(\Omega), R_h\varphi \in \mathcal{V}_h(0): \quad \langle \nabla(R_h\varphi - \varphi), \nabla\chi\rangle = 0 \quad \forall\, \chi \in \mathcal{V}_h(0). \end{aligned} \tag{4.7}$$

LEMMA 4.4. *The following quasi-optimal error estimates are valid for R_h* [35]*:*

$$\|\varphi - R_h\varphi\|_{L^p(\Omega)} \le Cph\|\nabla(\varphi - \Pi_h\varphi)\|_{L^p(\Omega)} \quad \forall\, p \ge 2, \tag{4.8}$$

$$\|\varphi - R_h\varphi\|_{L^\infty(\Omega)} \le Ch|\log h|\|\nabla(\varphi - \Pi_h\varphi)\|_{L^\infty(\Omega)}. \tag{4.9}$$

In view of (4.7), and the notation for errors

$$\rho^n := Y^n - R_h(\bar{y}^n), \quad \eta^n := y^n - R_h(\bar{y}^n),$$

we are able to rewrite (4.4) as follows:

$$\langle \rho^n - \rho^{n-1}, \chi\rangle + \tau\langle \nabla\rho^n, \nabla\chi\rangle = \langle \eta^n - \eta^{n-1}, \chi\rangle \quad \forall\ \chi \in \mathcal{V}_h(0). \tag{4.10}$$

This *error equation* suggests considering the following (backward) *dual discrete parabolic problem* with terminal datum $Z^m := \zeta \in \mathcal{V}_h(0)$ for $1 \le m \le N$ [11], [24]: for any $1 \le n \le m$ find $Z^{n-1} \in \mathcal{V}_h(0)$ satisfying

$$\langle Z^{n-1} - Z^n, \chi\rangle + \tau\langle \nabla Z^{n-1}, \nabla\chi\rangle = 0 \quad \forall\ \chi \in \mathcal{V}_h(0). \tag{4.11}$$

If we now take $\chi = Z^{n-1} \in \mathcal{V}_h(0)$ in (4.10), sum by parts from $n = 1$ to $n = m$, and use (4.11) together with the fact that $\rho^n \in \mathcal{V}_h(0)$, we easily arrive at

$$\langle \rho^m, \zeta\rangle = \langle \eta^m, \zeta\rangle + \sum_{n=1}^{m}\langle \eta^n, Z^{n-1} - Z^n\rangle + \langle Y^0 - y^0, Z^0\rangle.$$

In order to get an L^∞ bound for ρ^m we choose now $\zeta := P_h(\sigma^{-2}\rho^m)$. Hence

$$\|\sigma^{-1}\rho^m\|^2_{L^2(\Omega)} = \langle \eta^m, P_h(\sigma^{-2}\rho^m)\rangle + \sum_{n=1}^{m}\langle \eta^n, Z^{n-1} - Z^n\rangle + \langle Y^0 - y^0, Z^0\rangle.$$

Since $\|\sigma^s P_h v\|_{L^2(\Omega)} \le C\|\sigma^s v\|_{L^2(\Omega)}$ for any $s \in \mathbb{R}$, we get

$$\begin{aligned}\langle \eta^m, P_h(\sigma^{-2}\rho^m)\rangle &\le \|\sigma^{-1}\eta^m\|_{L^2(\Omega)}\|\sigma P_h(\sigma^{-2}\rho^m)\|_{L^2(\Omega)}\\ &\le C\|\sigma^{-1}\eta^m\|_{L^2(\Omega)}\|\sigma^{-1}\rho^m\|_{L^2(\Omega)},\end{aligned}$$

whence

$$\begin{aligned}\|\sigma^{-1}\rho^m\|^2_{L^2(\Omega)} \le C\|\sigma^{-1}\eta^m\|^2_{L^2(\Omega)} &+ 2\sum_{n=1}^{m}\|\sigma^{-1}\eta^n\|_{L^2(\Omega)}\|\sigma(Z^n - Z^{n-1})\|_{L^2(\Omega)}\\ &+ 2\|\sigma^{-1}(\Pi_h y^0 - y^0)\|_{L^2(\Omega)}\|\sigma Z^0\|_{L^2(\Omega)}.\end{aligned} \tag{4.12}$$

It is thus evident that we need a priori estimates for $\{Z^n - Z^{n-1}\}_{n=1}^N$ and Z^0 in weighted L^2-based norms. We simply state two key results from [11, eqq. (5.3),(5.4)], which are valid even for mildly graded meshes and variable time steps.

LEMMA 4.5. *There exists a constant $C^* \ge 1$ such that, if $\omega = C^*h|\log h|$ in (4.5) and $\tau \ge C^*h^2|\log h|^3$, then* [11]

$$\|\sigma Z^0\|_{L^2(\Omega)} + \Big(\tau^{-1}\sum_{n=1}^{m} t_{m-n+1}\|\sigma(Z^n - Z^{n-1})\|^2_{L^2(\Omega)}\Big)^{1/2} \le C|\log h|\|\sigma\zeta\|_{L^2(\Omega)}.$$

Since $\|\sigma\zeta\|_{L^2(\Omega)} \le C\|\sigma^{-1}\rho^m\|_{L^2(\Omega)}$, Lemma 4.5 and (4.12) must be supplied with the following asymptotic weighted norm estimates for $\{\eta^n\}_{n=1}^N$ and $\Pi_h y^0 - y^0$ as $p \uparrow \infty$.

LEMMA 4.6.

$$\Big(\tau\sum_{n=1}^{m} t^{-1}_{m-n+1}\|\sigma^{-1}\eta^n\|^2_{L^2(\Omega)}\Big)^{1/2} \le Cph^{-2/p}\tau^{-1/p}\Big(\tau\sum_{n=1}^{m}\|\eta^n\|^p_{L^p(\Omega)}\Big)^{1/p}.$$

Proof. By Hölder inequality, with $q = p/(p-2)$, we infer that

$$\tau \sum_{n=1}^{m} t_{m-n+1}^{-1} \int_\Omega \sigma^{-2} |\eta^n|^2 \le \Big(\tau \sum_{n=1}^{m} t_{m-n+1}^{-q} \int_\Omega \sigma^{-2q}\Big)^{1/q} \Big(\tau \sum_{n=1}^{m} \|\eta^n\|_{L^p(\Omega)}^p\Big)^{2/p}.$$

So we just have to deal with the first term on the right hand side. By virtue of Lemma 4.3 with $s = 2q > 2$, and the fact that $\omega \ge h$, as $p \uparrow \infty$ we end up with

$$\Big(\tau \sum_{n=1}^{m} t_{m-n+1}^{-q} \int_\Omega \sigma^{-2q}\Big)^{1/q} \le C\Big(\int_\tau^\infty t^{-q} dt \int_\Omega \sigma^{-2q}\Big)^{1/q} \le Cp^2 h^{-4/p} \tau^{-2/p}. \qquad \square$$

LEMMA 4.7. $\|\sigma^{-1}\eta^m\|_{L^2(\Omega)} \le Cph^{-2/p}\tau^{-1/p}\Big(\tau \sum_{n=1}^{m} \|\eta^n\|_{L^p(\Omega)}^p\Big)^{1/p}.$

Proof. Let $\xi : [0,T] \to \mathbb{R}$ be the auxiliary weight function $\xi(t) := \big(|t - t_m|^2 + \tau^2\big)^{1/2}$. If $q = p/(p-2)$, then by Hölder inequality

$$\begin{aligned}\int_\Omega \sigma^{-2}|\eta^m|^2 &\le \tau \sum_{n=1}^{m} (\xi^n)^{-1} \int_\Omega \sigma^{-2}|\eta^n|^2 \\ &\le \Big(\tau \sum_{n=1}^{m} (\xi^n)^{-q} \int_\Omega \sigma^{-2q}\Big)^{1/q} \Big(\tau \sum_{n=1}^{m} \|\eta^n\|_{L^p(\Omega)}^p\Big)^{2/p}.\end{aligned}$$

The assertion thus follows as in Lemma 4.6 from $\tau \sum_{n=1}^{m} (\xi^n)^{-q} \le C \int_\tau^\infty t^{-q} dt$. $\square$

LEMMA 4.8. $\|\sigma^{-1}(\Pi_h y^0 - y^0)\|_{L^2(\Omega)} \le Cp^{1/2}h^{2-4/p}\|y\|_{W_p^{2,1}(Q)}$.

Proof. Apply Hölder inequality, with $q = p/(p-2)$, to get

$$\|\sigma^{-1}(\Pi_h y^0 - y^0)\|_{L^2(\Omega)} \le \Big(\int_\Omega \sigma^{-2q}\Big)^{1/2q} \|\Pi_h y^0 - y^0\|_{L^p(\Omega)}.$$

Then use Lemma 4.3 with $s = 2q$, the trace inequality [21, p.342]

$$\|y^0\|_{W_p^{2-2/p}(\Omega)} \le C\|y\|_{W_p^{2,1}(Q)},$$

and (4.6) for $r = p$ and $s = 2 - 2/p$, to deduce the desired result as $p \uparrow \infty$. $\square$

In order to complete the argument we must get the following bound as $p \uparrow \infty$.

LEMMA 4.9. $\Big(\tau \sum_{n=1}^{m} \|\eta^n\|_{L^p(\Omega)}^p\Big)^{1/p} \le Cp(h^2 + \tau)\|y\|_{W_p^{2,1}(Q)}$.

Proof. We first rewrite η^n as follows:

$$\eta^n = (y^n - \bar{y}^n) + (\bar{y}^n - R_h(\bar{y}^n)) = \tfrac{1}{\tau} \int_{t_{n-1}}^{t_n} \big(y(t_n) - y(t)\big) dt + \tfrac{1}{\tau}(I - R_h) \int_{t_{n-1}}^{t_n} y(t) dt.$$

Applying Hölder inequality, combined with (4.8) and (4.6) for $r = p$ and $s = 2$, we get

$$\begin{aligned}\Big(\tau \sum_{n=1}^{m} \|\eta^n\|_{L^p(\Omega)}^p\Big)^{1/p} &\le C\tau\Big(\sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \|\partial_t y\|_{L^p(\Omega)}^p\Big)^{1/p} \\ &\quad + Cph^2\Big(\sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \|D^2 y\|_{L^p(\Omega)}^p\Big)^{1/p},\end{aligned}$$

which yields the desired estimate. $\square$

Proof of Theorem 4.1. First note that, for all $t_{m-1} < t \le t_m$, $1 \le m \le N$,

$$y(\cdot,t) - \hat{H}_{h,\tau}(y;\cdot,t) = \tfrac{t-t_{m-1}}{\tau}\big(y(\cdot,t) - Y^m(\cdot)\big) + \tfrac{t_m - t}{\tau}\big(y(\cdot,t) - Y^{m-1}(\cdot)\big).$$

Since $\rho^m = Y^m - R_h(\bar{y}^m)$ is piecewise linear, an inverse inequality combined with (4.5) for $\mathbf{x}_0 \in \Omega$ being the point at which ρ^m attains the maximum norm [22], yields

$$\|y(t) - Y^m\|_{L^\infty(\Omega)} \le \|y(t) - R_h(\bar{y}^m)\|_{L^\infty(\Omega)} + \tfrac{\omega}{h}\|\sigma^{-1}\rho^m\|_{L^2(\Omega)} \quad \forall\ t_{m-1} < t \le t_m.$$

In view of (4.12), Lemmas 4.5 to 4.9, and the fact that $\omega = Ch|\log h|$, we infer that

$$\tfrac{\omega}{h}\|\sigma^{-1}\rho^m\|_{L^2(\Omega)} \le Cp^2|\log h|^2\big(h^{2-4/p} + \tau^{1-1/p}\big)\|y\|_{W_p^{2,1}(Q)}.$$

We then resort to (3.2), (4.9), and (4.6) for $r = \infty$ and $s = 2 - 4/p$, to get

$$\begin{aligned}\|y(t) - R_h(\bar{y}^m)\|_{L^\infty(\Omega)} &\le \|y(t) - \bar{y}^m\|_{L^\infty(\Omega)} + \|\bar{y}^m - R_h(\bar{y}^m)\|_{L^\infty(\Omega)} \\ &\le \tfrac{1}{\tau}\int_{t_{m-1}}^{t_m}\big(\|y(t) - y(s)\|_{L^\infty(\Omega)} + Ch^{2-4/p}|\log h|\|y(s)\|_{W_\infty^{2-4/p}(\Omega)}\big)ds \\ &\le C\big(\tau^{1-2/p} + h^{2-4/p}|\log h|\big)\|y\|_{W_p^{2,1}(Q)}.\end{aligned}$$

Likewise we can estimate $\|y(t) - Y^{m-1}\|_{L^\infty(\Omega)}$ and this concludes the proof. ∎

Let $H_{h,\tau}$ denote the parabolic projection associated with the heat operator (4.4) and Dirichlet boundary condition 1: $H_{h,\tau}(1+y) := 1 + \hat{H}_{h,\tau}(y)$. The following crucial estimate, proven here in 2D, will be assumed valid in the sequel for $n \ge 2$.

COROLLARY 4.1. *If u is the solution to* (3.1) *with forcing term of order $\mathcal{O}(\varepsilon^{-1})$, then*

$$\|u - H_{h,\tau}(u)\|_{L^\infty(Q)} \le C\tfrac{h^2+\tau}{\varepsilon^2}|\log h|^5 \quad \textit{for } \tau \ge C^* h^2 |\log h|^3.$$

Proof. Take $y = u - 1$ in Theorem 4.1 and use Lemma 3.1 with $p = |\log h|$. ∎

5. Fully discrete barriers. We now intend to construct a fully discrete supersolution $V^+_{\varepsilon,h,\tau}$ to (4.2). Let u^+ be the solution to (3.1) with perturbed forcing term $g + \alpha$ and $\phi_\varepsilon = \gamma(d_0/\varepsilon)$, and $V^+_{\varepsilon,h,\tau}$ be thereby defined as follows for $\Lambda \ge 1$ to be chosen later:

$$V^+_{\varepsilon,h,\tau} := H_{h,\tau}(u^+) + \Lambda\,\xi(\varepsilon,h,\tau), \quad \xi(\varepsilon,h,\tau) := \|u^+ - H_{h,\tau}(u^+)\|_{L^\infty(Q)}. \tag{5.1}$$

In view of Corollary 4.1, $\xi(\varepsilon,h,\tau) \le C(h^2+\tau)\varepsilon^{-2}|\log h|^5$ for $\tau \ge C^* h^2|\log h|^3$, which asymptotically supersedes $\tau \ge C_\theta^{-1}h^2$ and will be assumed from now on.

LEMMA 5.1. *If Λ is sufficiently large and $\alpha \ge 4\Lambda\xi/c_0\varepsilon$, then $V^+_{\varepsilon,h,\tau}$ is a supersolution of* (4.2)*, namely it satisfies $V^+_{\varepsilon,h,\tau} \ge -1$ and, for all nodes $\mathbf{x}_i$ such that $V^{+,n}_{\varepsilon,h,\tau}(\mathbf{x}_i) < 1$,*

$$\varepsilon\langle\partial V^{+,n}_{\varepsilon,h,\tau}, \phi_i\rangle + \varepsilon\langle\nabla V^{+,n}_{\varepsilon,h,\tau}, \nabla\phi_i\rangle - \tfrac{1}{\varepsilon}\langle V^{+,n}_{\varepsilon,h,\tau}, \phi_i\rangle \ge \tfrac{c_0}{2}\langle g^n, \phi_i\rangle. \tag{5.2}$$

Proof. Since no confusion is possible we set $u = u^+$, $V = V^+_{\varepsilon,h,\tau}$, $H = H_{h,\tau}(u)$, and $\xi = \xi(\varepsilon,h,\tau)$. In view of Corollary 4.1 and (5.1), we see that, for all $(\mathbf{x},t) \in Q$,

$$-1 \le u \le H + \xi = V - (\Lambda - 1)\xi \le V \quad \forall\ \Lambda \ge 1. \tag{5.3}$$

If $V^n(\mathbf{x}_i) < 1$, then (5.3) implies $u^n(\mathbf{x}_i) < 1$. We claim that, upon properly choosing Λ, there exists $\mu_i^n < 1$ such that $u(\mathbf{x},t) \le \mu_i^n$ for all points $(\mathbf{x},t) \in K_i^n$. If not, there must exist $(\mathbf{x},t) \in K_i^n$ such that $u(\mathbf{x},t)=1$; hence $\nabla u(\mathbf{x},t)=0$ because $u \le 1$. We can use Lemma 3.2, for $y=u$, and Lemma 3.1 to arrive at

$$|\nabla u(\mathbf{y},t)| = |\nabla u(\mathbf{x},t) - \nabla u(\mathbf{y},t)| \le \omega(|\mathbf{x}-\mathbf{y}|) \le C\tfrac{h|\log h|}{\varepsilon^2}$$

for all $\mathbf{y} = s\mathbf{x}_i + (1-s)\mathbf{x}$ $(0<s<1)$. Applying the Mean Value Theorem we get

$$u(\mathbf{x}_i,t) = u(\mathbf{x},t) + \int_0^1 \big(\nabla u(s\mathbf{x}_i + (1-s)\mathbf{x},t)\big) ds \cdot (\mathbf{x}_i - \mathbf{x}) \ge 1 - C\tfrac{h^2|\log h|}{\varepsilon^2}.$$

This, combined again with Lemmas 3.1 and 3.2 for estimating $u(\mathbf{x}_i,t) - u(\mathbf{x}_i,t_n)$, yields

$$u(\mathbf{x}_i,t_n) \ge 1 - C\tfrac{h^2|\log h|}{\varepsilon^2} - C\tfrac{\tau|\log\tau|}{\varepsilon^2},$$

which is a contradiction with (5.3) at $(\mathbf{x}_i,t_n)$ provided Λ is sufficiently large.

The above claim enables us to take $\varphi = u + (1-\mu_i^n)\phi_i \in \mathcal{K}$ in (3.1), with g replaced by $g+\alpha$, to obtain

$$\varepsilon\langle \partial_t u, \phi_i\rangle + \varepsilon\langle \nabla u, \nabla\phi_i\rangle - \tfrac{1}{\varepsilon}\langle u,\phi_i\rangle \ge \tfrac{c_0}{2}\langle g+\alpha,\phi_i\rangle \quad \text{for a.e. } t_{n-1}<t<t_n,$$

whence, integrating on (t_{n-1},t_n) and making use of (4.4) with $y=u-1$, we get

$$\varepsilon\langle \partial H^n, \phi_i\rangle + \varepsilon\langle \nabla H^n, \nabla\phi_i\rangle \ge \tfrac{1}{\varepsilon}\langle \bar{u}^n,\phi_i\rangle + \tfrac{c_0}{2}\langle \bar{g}^n+\alpha,\phi_i\rangle.$$

Applying Corollary 4.1 we find that

$$\begin{aligned}
\varepsilon\langle \partial V^n, \phi_i\rangle &+ \varepsilon\langle \nabla V^n, \nabla\phi_i\rangle - \tfrac{1}{\varepsilon}\langle V^n,\phi_i\rangle \\
&\ge \langle \tfrac{1}{\varepsilon}(\bar{u}^n - V^n) + \tfrac{c_0}{2}(\bar{g}^n - g^n) + \tfrac{c_0}{2}\alpha + \tfrac{c_0}{2}g^n, \phi_i\rangle \\
&\ge \langle -2\Lambda\tfrac{\xi}{\varepsilon} + \tfrac{c_0}{2}\alpha + \tfrac{c_0}{2}g^n,\phi_i\rangle \ge \tfrac{c_0}{2}\langle g^n,\phi_i\rangle,
\end{aligned}$$

provided $\alpha \ge 4\Lambda\xi/c_0\varepsilon = C(h^2+\tau)\varepsilon^{-3}|\log h|^5$. This implies the assertion. ∎

It is important to realize the essential difference between the above proof and that in [14]. That is the role of the perturbation α that compensates for the antimonotone behavior of the potential $\psi(s) = -s$, which is absent in the parabolic obstacle problem of [14] as corresponds to the one-phase Stefan problem.

We denote by $U_{\varepsilon,h,\tau}$ the continuous piecewise linear function in space and time determined from $\{U^n\}_{n=0}^N$ by

$$U_{\varepsilon,h,\tau}(\cdot,t) := \tfrac{t-t_{n-1}}{\tau}U^n(\cdot) + \tfrac{t_n-t}{\tau}U^{n-1}(\cdot) \quad \forall\, t_{n-1}<t\le t_n,\ 1\le n\le N.$$

LEMMA 5.2. *Under the assumptions of Lemma 5.1 it holds* $U_{\varepsilon,h,\tau} \le V^+_{\varepsilon,h,\tau}$.

Proof. For simplicity we set $U = U_{\varepsilon,h,\tau}, V = V^+_{\varepsilon,h,\tau}$ and $S = U - V$. The above statement is equivalent to showing that $[S^n]^+ := \Pi_h\big[\max(S^n,0)\big] \in \mathcal{V}_h(0)$ is identically zero for all $0 \le n \le N$. This is certainly true for $n=0$, because $V^0 = U^0 + \Lambda\xi$. Assume then by induction $S^{n-1}\le 0$. Note that $[S^n]^+(\mathbf{x}_i) > 0$ means $1 \ge U^n(\mathbf{x}_i) > V^n(\mathbf{x}_i)$. This, together with (5.2) multiplied by $\max(S^n(\mathbf{x}_i),0)$, yields

$$\varepsilon\langle \partial V^n, [S^n]^+\rangle + \varepsilon\langle \nabla V^n, \nabla[S^n]^+\rangle - \tfrac{1}{\varepsilon}\langle V^n,[S^n]^+\rangle \ge \tfrac{c_0}{2}\langle g^n,[S^n]^+\rangle. \tag{5.4}$$

On the other hand, we can take $\chi = \Pi_h\big[\min(U^n,V^n)\big] \in \mathcal{K}_h$ in (4.2) to obtain

$$\varepsilon\langle \partial U^n, [S^n]^+\rangle + \varepsilon\langle \nabla U^n, \nabla[S^n]^+\rangle - \tfrac{1}{\varepsilon}\langle U^n,[S^n]^+\rangle \le \tfrac{c_0}{2}\langle g^n,[S^n]^+\rangle, \tag{5.5}$$

because $\chi - U^n = -[S^n]^+$. Subtracting (5.4) from (5.5), and reordering, we get

$$\big(\tfrac{1}{\tau} - \tfrac{1}{\varepsilon^2}\big)\langle S^n,[S^n]^+\rangle + \langle \nabla S^n, \nabla[S^n]^+\rangle \le \tfrac{1}{\tau}\langle S^{n-1},[S^n]^+\rangle \le 0,$$

by induction assumption. Then on invoking Lemma 4.1 we deduce that for $\tau \geq C_\theta^{-1} h^2$ the above inequality results in $[S^n]^+ = 0$, as desired. $\square$

In the same spirit of (5.1) we can introduce the fully discrete subsolution

$$V^-_{\varepsilon,h,\tau} := H_{h,\tau}(u^-) - \Lambda\, \xi(\varepsilon, h, \tau) \leq U_{\varepsilon,h,\tau}.$$

As already discussed after Lemma 3.2, it is not feasible to derive a meaningful lower or upper bound for $U_{\varepsilon,h,\tau}$ only in terms of $V^+_{\varepsilon,h,\tau}$ or $V^-_{\varepsilon,h,\tau}$. We can think of both $V^+_{\varepsilon,h,\tau}(\cdot,t)$ and $V^-_{\varepsilon,h,\tau}(\cdot,t)$ as suitable discrete horizontal shifts of $U_{\varepsilon,h,\tau}(\cdot,t)$ in the direction normal to $\Sigma_{\varepsilon,h,\tau}(t)$ but in opposite senses.

6. Discrete nondegeneracy. In order to prove that $U_{\varepsilon,h,\tau}$ satisfies a nondegeneracy property, we need some further notation. For any $(\mathbf{x}, t) \in Q$ and $r > 0$, we consider the cylinder

$$Q_r(\mathbf{x}, t) := \{(\mathbf{y}, s) \in Q : |\mathbf{y} - \mathbf{x}| < r, 0 < t - s < r^2\}$$

and its discrete counterpart $Q_r^{h,\tau}(\mathbf{x}, t)$ of all sets K_i^n contained in $\bar{Q}_r(\mathbf{x}, t)$. For $r \geq A(h^2 + \tau)^{1/2}$ with $A \geq 1$ sufficiently large and h, τ small, $Q_r^{h,\tau}(\mathbf{x}, t)$ is nonempty. We denote by $\partial_p Q_r^{h,\tau}(\mathbf{x}, t)$ the parabolic boundary of $Q_r^{h,\tau}(\mathbf{x}, t)$, and note that the parabolic distance defined by $\operatorname{dist}_p\big((\mathbf{x}, t), (\mathbf{y}, s)\big) := \max\big(|\mathbf{x} - \mathbf{y}|, |t - s|^{1/2}\big)$ satisfies

$$\operatorname{dist}_p\big((\mathbf{x}, t), (\mathbf{y}, s)\big) \geq r - C(h + \tau^{1/2}) \geq Cr \quad \forall\ (\mathbf{y}, s) \in \partial_p Q_r^{h,\tau}(\mathbf{x}, t) \backslash \partial_p Q. \tag{6.1}$$

LEMMA 6.1. *There exist $\varepsilon^*, h^*, \tau^* > 0$ such that for all $h \leq h^*$, $\tau^* \geq \tau \geq C^* h^2 |\log h|^3$, $\varepsilon^* \geq \varepsilon \geq r \geq A(h^2 + \tau)^{1/2}$, and any node $(\mathbf{x}_j, t_m)$ satisfying $-1 < U^m(\mathbf{x}_j) \leq -1/2$ (resp. $1/2 \leq U^m(\mathbf{x}_j) < 1$) and either $t_m \geq r^2$ or $U^0(\mathbf{x}_i) = -1$ (resp. $U^0(\mathbf{x}_i) = 1$) for all nodes $\mathbf{x}_i$ such that $|\mathbf{x}_i - \mathbf{x}_j| < r$, there exists a node $(\mathbf{x}_i, t_n) \in Q_j^m := Q_r^{h,\tau}(\mathbf{x}_j, t_m)$ such that*

$$U^n(\mathbf{x}_i) \geq U^m(\mathbf{x}_j) + C\tfrac{r^2}{\varepsilon^2} \quad (\textit{resp. } U^n(\mathbf{x}_i) \leq U^m(\mathbf{x}_j) - C\tfrac{r^2}{\varepsilon^2}). \tag{6.2}$$

Proof. We only consider the case $-1 < U^m(\mathbf{x}_j) \leq -1/2$, and assume $U^n(\mathbf{x}_i) \leq -1/4$ for all nodes $(\mathbf{x}_i, t_n) \in Q_j^m$ because otherwise (6.2) is trivial. For all those nodes, we can take $\chi = U^n - (1 + U^n(\mathbf{x}_i))\phi_i$ in (4.2) to obtain, for ε sufficiently small,

$$\langle \partial U^n, \phi_i\rangle + \langle \nabla U^n, \nabla \phi_i\rangle \leq \tfrac{1}{\varepsilon^2}\langle U^n + \varepsilon \tfrac{c_0}{2} g^n, \phi_i\rangle \leq -\tfrac{1}{8\varepsilon^2}\langle 1, \phi_i\rangle, \tag{6.3}$$

provided $U^n(\mathbf{x}_i) > -1$. Let $D \subset Q_j^m$ be the (closed) set containing $(\mathbf{x}_j, t_m)$ which is the connected union of all those K_i^n's of Q_j^m such that $U^n(\mathbf{x}_i) > -1$. The parabolic boundary $\partial_p D$ of D is made of those nodes $(\mathbf{x}_i, t_n) \in D$ such that $K_i^n \not\subset D$ or $(\mathbf{x}_i, t_n) \in \partial_p Q$, the latter reducing to $t_n = 0$ because $\operatorname{dist}(\mathbf{x}_j, \partial\Omega) > r$. Consequently $(\mathbf{x}_i, t_n)$ satisfies either (a) $U^n(\mathbf{x}_i) > -1$ and $(\mathbf{x}_i, t_n) \in \partial_p Q_j^m \backslash \partial_p Q$ or (b) $U^n(\mathbf{x}_i) = -1$, which means that $(\mathbf{x}_i, t_n)$ belongs either to the discrete free boundary $\partial\{U^n > -1\}$ or the initial coincidence set $\{U^0 = -1\}$.

Let $\sigma(\cdot, t) \in \mathcal{V}_h$ be given by

$$\sigma(\mathbf{x}, t) := \Pi_h\big(|\mathbf{x} - \mathbf{x}_j|^2\big) + t_m - t.$$

Note that σ satisfies, for a constant $\beta > 0$ independent of h, τ, and $(\mathbf{x}_j, t_m)$ [5], [33],

$$\langle \partial\sigma^n, \phi_i\rangle + \langle \nabla\sigma^n, \nabla\phi_i\rangle \geq -\beta\langle 1, \phi_i\rangle \quad \forall\ 1 \leq i \leq I,\ 1 \leq n \leq N. \tag{6.4}$$

Combining (6.3) with (6.4), the functions $\{S^n := U^n - \sigma^n/8\beta\varepsilon^2\}_{n=0}^N$ satisfy

$$\langle \partial S^n, \phi_i\rangle + \langle \nabla S^n, \nabla\phi_i\rangle \leq 0 \quad \forall\ (\mathbf{x}_i, t_n) \in D\backslash\partial_p D.$$

Since $\tau \geq C_\theta^{-1} h^2$, Lemma 4.2 yields the existence of $(\mathbf{x}_i, t_n) \in \partial_p D$ such that $S^n(\mathbf{x}_i) \geq S^m(\mathbf{x}_j) = U^m(\mathbf{x}_j) > -1$, and so

$$U^n(\mathbf{x}_i) - U^m(\mathbf{x}_j) \geq \tfrac{1}{8\beta\varepsilon^2}\sigma^n(\mathbf{x}_i).$$

Since condition (b) above, namely $U^n(\mathbf{x}_i) = -1$, leads to the contradiction $S^n(\mathbf{x}_i) = -1 - \sigma^n(\mathbf{x}_i)/8\beta\varepsilon^2 < -1$, then (a) applies and (6.1) thus implies

$$\sigma^n(\mathbf{x}_i) = \sigma(\mathbf{x}_i, t_n) = |\mathbf{x}_i - \mathbf{x}_j|^2 + t_m - t_n = \operatorname{dist}_p^2\big((\mathbf{x}_i, t_n), (\mathbf{x}_j, t_m)\big) \geq Cr^2.$$

Combining the last two inequalities we obtain (6.2) as desired. The case $1/2 \leq U^m(\mathbf{x}_j) < 1$ is similar except that we cannot longer use $\operatorname{dist}(\mathbf{x}_j, \partial\Omega) > r$ but instead $U^n = 1$ on $\partial\Omega$. □

In view of (6.2) we realize that the discrete solution $U_{\varepsilon,h,\tau}$ detaches from the obstacles -1 and 1 with a prescribed quadratic growth, which is a weak nondegeneracy statement. A similar result can be derived for u_ε as well.

7. Convergence. Let $\Sigma_{\varepsilon,h,\tau}$ indicate the discrete zero level set and $\mathcal{T}_{\varepsilon,h,\tau}$ the discrete transition region (or noncoincidence set), namely

$$\Sigma_{\varepsilon,h,\tau}(t) := \{\mathbf{x} \in \Omega : U_{\varepsilon,h,\tau}(\mathbf{x}, t) = 0\}, \quad \mathcal{T}_{\varepsilon,h,\tau}(t) := \big\{\mathbf{x} \in \Omega : |U_{\varepsilon,h,\tau}(\mathbf{x}, t)| < 1\big\}.$$

In this section we prove that, as $\varepsilon \downarrow 0$, both $\Sigma_{\varepsilon,h,\tau}(t)$ and $\mathcal{T}_{\varepsilon,h,\tau}(t)$ converge to $\Sigma(t)$, provided $\Sigma(t)$ does not develop interior (no fattening). This reproduces for the fully discrete problem the convergence result of [30].

For $r > 0$ consider the parabolic ball $B_r(\mathbf{x}, t) := \big\{(\mathbf{y}, s) \in Q : \operatorname{dist}_p\big((\mathbf{x}, t), (\mathbf{y}, s)\big) < r\big\}$. Hereafter we employ the notation of §3, so w and $w^\pm$ stand for viscosity solutions to (1.2) with forcing terms g and $g \pm \alpha$ and initial traces d_0 and $d_0 \pm l$, respectively. Consider $\mathbf{x} \in I(t)$, the inside of $\Sigma(t)$, that is $w(\mathbf{x}, t) < 0$. As a consequence of continuous dependence of viscosity solutions we infer that $w^+(\mathbf{x}, t) < 0$ for $\alpha > 0$ sufficiently small depending on $(\mathbf{x}, t)$. Since w^+ is continuous in space and time, there exists $r > 0$ sufficiently small such that

$$w^+(\mathbf{y}, s) < 0 \quad \forall\ (\mathbf{y}, s) \in B_{2r}(\mathbf{x}, t).$$

Now we use the same argument of [30, Theorem 7.1]. Since $d^+(\cdot, t)$ is the signed distance to $\Sigma^+(t)$, the zero level set of $w^+(\cdot, t)$, it turns out that $d^+ < 0$ in $B_{2r}(\mathbf{x}, t)$, whence

$$\iota^+(\mathbf{y}, s) = \tfrac{d^+(\mathbf{y}, s)}{\varepsilon} - \sigma(t) < -\sigma(t) < -\tfrac{\pi}{2} \quad \forall\ (\mathbf{y}, s) \in B_{2r}(\mathbf{x}, t). \tag{7.1}$$

If $\Sigma^+(t)$ develops interior, then $d^+(\mathbf{x}, t)$ is replaced by $\operatorname{dist}(\mathbf{x}, \partial\{w^+(\cdot, t) < 0\})$ and the reasoning still applies. Invoking Lemma 3.3, for $\alpha \geq \varepsilon$ we deduce

$$-1 \leq u^+(\mathbf{y}, s) \leq v^+(\mathbf{y}, s) = -1 \quad \forall\ (\mathbf{y}, s) \in B_{2r}(\mathbf{x}, t).$$

If Λ is sufficiently large and $\alpha \geq 4\Lambda\xi/c_0\varepsilon$, then Lemma 5.2 together with (5.1) and Corollary 4.1 implies

$$\begin{aligned} U_{\varepsilon,h,\tau}(\mathbf{y}, s) &\leq V^+_{\varepsilon,h,\tau}(\mathbf{y}, s) \leq u^+(\mathbf{y}, s) + (\Lambda + 1)\xi(\varepsilon, h, \tau) \\ &\leq -1 + C\tfrac{h^2+\tau}{\varepsilon^2}|\log h|^5 \leq -\tfrac{1}{2} \quad \forall\ (\mathbf{y}, s) \in B_{2r}(\mathbf{x}, t). \end{aligned} \tag{7.2}$$

Suppose by contradiction that $U_{\varepsilon,h,\tau}(\mathbf{x}, t) > -1$. For ε, h, and τ sufficiently small satisfying $\varepsilon \geq r \geq A(h^2+\tau)^{1/2}$, there thus exists an adjacent node $(\mathbf{x}_j, t_m) \in B_{2r}(\mathbf{x}, t)$

where $-1 < U_{\varepsilon,h,\tau}(\mathbf{x}_j, t_m) = U^m(\mathbf{x}_j) \le -1/2$ and $t_m > t - \tau \ge r^2$. We can then apply Lemma 6.1 to obtain another node $(\mathbf{x}_i, t_n)$ so that

$$(7.3) \qquad (\mathbf{x}_i, t_n) \in Q_r^{h,\tau}(\mathbf{x}_j, t_m) \subset B_{2r}(\mathbf{x}, t), \quad U_{\varepsilon,h,\tau}(\mathbf{x}_i, t_n) > -1 + C\tfrac{r^2}{\varepsilon^2}.$$

On selecting $r = R(h^2 + \tau)^{1/2} |\log h|^{5/2}$ for a sufficiently large R, we find that (7.2) and (7.3) cannot hold at the same time; hence $U_{\varepsilon,h,\tau}(\mathbf{x}, t) = -1$. Likewise, if $\mathbf{x} \in O(t)$ then $U_{\varepsilon,h,\tau}(\mathbf{x}, t) = 1$ asymptotically.

In addition to the condition $\alpha \le \alpha^*(\mathbf{x}, t) = o(1)$ used above, we recall the constraints $\alpha \ge \varepsilon$, $\alpha \ge 4\Lambda\xi/c_0\varepsilon = C(h^2+\tau)\varepsilon^{-3}|\log h|^5$, and $\tau \ge C^* h^2 |\log h|^3$ for Lemmas 3.3 and 5.2 and Corollary 4.1 to hold. They in turn are compatible with the choice of r in that $A(h^2+\tau)^{1/2} \le r \le \varepsilon$ can be enforced asymptotically. The following corollary thus summarizes the above derivation.

Corollary 7.1. *Let $\mathbf{x} \in I(t)$ (resp. $O(t)$) and let $h = o(\varepsilon^{3/2}|\log \varepsilon|^{-4})$ and $\tau = C^* h^2 |\log h|^3$. Then there exists $\varepsilon^* > 0$ depending on $(\mathbf{x}, t)$ such that*

$$U_{\varepsilon,h,\tau}(\mathbf{x}, t) = -1 \quad (\textit{resp. } U_{\varepsilon,h,\tau}(\mathbf{x}, t) = 1) \quad \forall\, \varepsilon \le \varepsilon^*.$$

This convergence result, valid even past singularities, illustrates the need for both $u^\pm$ and $v^\pm$ in Lemma 3.3. The former is used in constructing the discrete barrier $V^\pm_{\varepsilon,h,\tau}$ because Lemma 3.1 is not necessarily valid for $v^\pm$, which may be discontinuous in time as $d^\pm$ is. Function $v^\pm$, however, provides a relation with the original front $\Sigma(t)$ because of continuous dependence of level sets. Without additional assumptions on $\Sigma(t)$, or the singularities, we cannot expect a rate of convergence. This issue is further explored in the next two sections.

8. Regular points.

We show, for regular points $(\mathbf{x}, t) \in \Sigma^*$, that both the distance between $\mathbf{x}$ and $\Sigma_{\varepsilon,h,\tau}(t)$ and the local thickness of $\mathcal{T}_{\varepsilon,h,\tau}(t)$ are $\mathcal{O}(\varepsilon)$. This last statement is remarkably important because it reveals that the local structure of the double obstacle formulation is retained by the fully discrete problem, a key computational goal. The occurrence of singularities, and related loss of error control, does not affect this linear rate of convergence, which happens to be valid irrespective of whether $(\mathbf{x}, t)$ lies before or in between singularities. This assertion is rather unexpected but in the same spirit of Corollary 2.2, that asserts Lipschitz dependence of interfaces under perturbations of data. Let $\mathrm{thick}(\mathcal{T}_{\varepsilon,h,\tau}(t); \mathbf{x}, \mathbf{n})$ denote the thickness of $\mathcal{T}_{\varepsilon,h,\tau}(t)$ in the direction $\mathbf{n}$ across $\mathbf{x}$.

Corollary 8.1. *Let $\Sigma^\sharp$ be a compact subset of Σ^*, and let $h = \mathcal{O}(\varepsilon^2|\log\varepsilon|^{-4})$ and $\tau = C^* h^2|\log h|^3$. Then there exist constants $\varepsilon^* > 0$ depending on $\Sigma^\sharp$ and $C > 0$ depending on data such that, for all $\varepsilon \le \varepsilon^*$ and all $(\mathbf{x}, t) \in \Sigma^\sharp$,*

$$\mathrm{dist}\big(\mathbf{x}, \Sigma_{\varepsilon,h,\tau}(t)\big),\ \mathrm{thick}\big(\mathcal{T}_{\varepsilon,h,\tau}(t); \mathbf{x}, \tfrac{\nabla w(\mathbf{x},t)}{|\nabla w(\mathbf{x},t)|}\big) \le \tfrac{C}{|\nabla w(\mathbf{x},t)|}\varepsilon.$$

Proof. For $\varepsilon, h, \tau > 0$ satisfying the above relations and $C \ge 1$ sufficiently large, let $\alpha = C\varepsilon$ satisfy $\alpha \ge 4\Lambda\xi/c_0\varepsilon$, so that Lemmas 3.3 and 5.2 and Corollary 4.1 are valid. Given $(\mathbf{x}, t) \in \Sigma^\sharp$, we argue along the lines of Corollary 2.2 and use the same notation. With $\mathbf{n} := \nabla w(\mathbf{x}, t)/|\nabla w(\mathbf{x}, t)|$, we readily have

$$w^+(\mathbf{x} - \delta\mathbf{n}, t) \le w(\mathbf{x} - \delta\mathbf{n}, t) + C\varepsilon \le -\tfrac{1}{2}\delta|\nabla w(\mathbf{x}, t)| + C\varepsilon = -\varepsilon,$$

provided $\delta = 2(C+1)\varepsilon/|\nabla w(\mathbf{x}, t)| \le D$, which holds for $\varepsilon \le \varepsilon^* := DW/2(C+1)$. Invoking Theorem 2.1 we can find a constant $0 < E \le 1$, only depending on data, such that $w^+ < 0$ in $B_{2r}(\mathbf{x} - \delta\mathbf{n}, t)$ is valid for all $r \le E\varepsilon$. Let $r = R(h^2+\tau)^{1/2}|\log h|^{5/2} \le E\varepsilon$ for $R \ge 1$ sufficiently large. Then we can use the argument of Corollary 7.1 to deduce $U_{\varepsilon,h,\tau}(\mathbf{x}-\delta\mathbf{n},t)=-1$. In fact, we either have $t_m \geq r^2$ or $t-\tau<t_m<r^2$, in which case (7.1) yields $\iota^+(\mathbf{x}_i,0)<-\pi/2$ for all nodes $|\mathbf{x}_i-(\mathbf{x}-\delta\mathbf{n})|<2r$, whence

$$-1 \leq U^0(\mathbf{x}_i)=\gamma\big(\tfrac{d_0(\mathbf{x}_i)}{\varepsilon}\big)=u^+(\mathbf{x}_i,0)\leq v^+(\mathbf{x}_i,0)=-1.$$

Consequently, Lemma 6.1 applies to both situations. The same reasoning can be repeated to obtain $U_{\varepsilon,h,\tau}(\mathbf{x}+\delta\mathbf{n},t)=1$. Since $U_{\varepsilon,h,\tau}$ is continuous in space there exists a point on the straight segment connecting $\mathbf{x}-\delta\mathbf{n}$ and $\mathbf{x}+\delta\mathbf{n}$ where $U_{\varepsilon,h,\tau}(\cdot,t)$ vanishes; hence $\operatorname{dist}\big(\mathbf{x},\Sigma_{\varepsilon,h,\tau}(t)\big)<\delta$. The thickness of $\mathcal{T}_{\varepsilon,h,\tau}(t)$ in the direction $\mathbf{n}$ passing through $\mathbf{x}$ can be finally estimated by 2δ. $\square$

We stress that the choice of ε, h, and τ satisfying $A(h^2+\tau)^{1/2}\leq r\leq E\varepsilon$ is not related here to the specific $(\mathbf{x},t)$ but rather to $\Sigma^\sharp$. This is an essential difference with Corollary 7.1 which reveals the crucial role played by Corollary 2.1.

9. Singularities. In this final section we elaborate on the space-time approximation of singularities under a further, but plausible, structural assumption. Let w be C^1 in a vicinity of a singular point $\mathbf{x}\in\Sigma(t)$, namely $\nabla w(\mathbf{x},t)=0$, and exhibit the following asymptotic behavior [16]: *there exists a constant β, a set of orthonormal vectors $\{\mathbf{e}_i\}_{i=1}^n$, and nonvanishing constants $\{\lambda_i\}_{i=1}^k$ with $1\leq k\leq n$ such that*

$$w(\mathbf{x}+\mathbf{y},t+s)=\beta s+\sum_{i=1}^k \lambda_i y_i^2+o\big(|s|+|\mathbf{y}|^2\big) \quad \text{for } \mathbf{y}=\sum_{i=1}^k y_i\mathbf{e}_i, \quad |\mathbf{y}|,|s|\ll 1. \tag{9.1}$$

This is certainly true if $w\in C^2$ in the vicinity of $(\mathbf{x},t)$, in which case $\{\lambda_i\}_{i=1}^k$ are the nontrivial eigenvalues of the Hessian $D^2w(\mathbf{x},t)$. However there is no characterization of C^2 regularity and, in fact, w may fail to be C^2 while still exhibiting quadratic growth in the directions $\{\mathbf{e}_i\}_{i=1}^k$ [20]. Substituting (9.1) into (1.2), it is not difficult to see that [16]

$$\lambda_i=\lambda\neq 0 \quad 1\leq i\leq k \quad \text{and} \quad \beta=2(k-1)\lambda+g(\mathbf{x},t),$$

which coupled with (9.1) yields the curvature blow-up rate of type I singularities [19]:

$$\kappa\approx|\mathbf{y}|^{-1}\approx\big|\tfrac{\lambda}{\beta s}\big|^{1/2} \quad \text{as } s\uparrow 0. \tag{9.2}$$

We now briefly discuss the particular but relevant case of the mean curvature flow, namely $g=0$, for which $\beta\neq 0$ provided $k\geq 2$. Convex interfaces would correspond to $k=n$ because they shrink asymptotically to round points [18], [19]. Likewise we can speculate that $2\leq k<n$ for surfaces of rotation such as those studied in [1], [25], [29], [36], in which case the singularity looks asymptotically like a cylinder. This would suggest that cylinders or spheres are the only possible asymptotic surfaces for type I singularities, which seems to be generic. Singularities of type II are those with blow-up rate higher than (9.2), and are receiving considerable attention. We conclude with the following result, whose proof proceeds as that of Corollary 8.1.

COROLLARY 9.1. *Let $\mathbf{x}\in\Sigma(t)$ satisfy (9.1) with $\beta,\lambda>0$. Let $h=\mathcal{O}(\varepsilon^2|\log\varepsilon|^{-4})$ and $\tau=C^*h^2|\log h|^3$. Then there exist constants $\varepsilon^*>0$ depending on $(\mathbf{x},t)$ and $C>0$ depending on data such that, for all $\varepsilon\leq\varepsilon^*$,*

$$U_{\varepsilon,h,\tau}(\mathbf{x}+\mathbf{y},t\pm s)=\pm 1 \quad \textit{for all } C\tfrac{\varepsilon}{\beta}\leq s\leq C\tfrac{\varepsilon^*}{\beta},\ |\mathbf{y}|^2<C\tfrac{\beta}{\lambda}s; \tag{9.3}$$

$$\textit{there exists } |s|\leq C\tfrac{\varepsilon}{\beta} \textit{ for which } \mathbf{x}\in\Sigma_{\varepsilon,h,\tau}(t+s); \tag{9.4}$$

$$\operatorname{thick}\big(\mathcal{T}_{\varepsilon,h,\tau}(t);\mathbf{x},\mathbf{e}_i\big)\leq C\big(\tfrac{\varepsilon}{\lambda}\big)^{1/2} \quad \textit{for all } 1\leq i\leq k. \tag{9.5}$$

Proof. Since Σ_0 is smooth, its flow $\Sigma(\cdot)$ is regular for a little while; thus $t > 0$. Given $\varepsilon, h, \tau > 0$ satisfying the above relations, let $\alpha = C\varepsilon$ satisfy $\alpha \geq 4\Lambda\xi/c_0\varepsilon$ for $C \geq 1$ sufficiently large. We use Corollary 2.1 in conjunction with (9.1) to arrive at

$$w^+(\mathbf{x}+\mathbf{y}, t-s) \leq w(\mathbf{x}+\mathbf{y}, t-s) + C\varepsilon = -\beta s + \lambda|\mathbf{y}|^2 + o(s) + C\varepsilon \leq -\frac{\beta}{4}s + C\varepsilon \leq -\varepsilon,$$

for $s \geq 4(C+1)\varepsilon/\beta$ and $|\mathbf{y}|^2 \leq \beta s/2\lambda$, provided s is sufficiently small so that $o(s) \leq \beta s/4$ and $t - s > 0$. Consequently, applying Theorem 2.1, there exists $E \leq 1$ such that $w^+ < 0$ in $B_{2r}(\mathbf{x}+\mathbf{y}, t-s)$ for $r = E\varepsilon$. Arguing as in Corollary 8.1 we infer that $U_{\varepsilon,h,\tau}(\mathbf{x}+\mathbf{y}, t-s) = -1$. We can similarly obtain $U_{\varepsilon,h,\tau}(\mathbf{x}+\mathbf{y}, t+s) = 1$, which yields (9.3) and so (9.4) because $U_{\varepsilon,h,\tau}$ is continuous in time. Finally, Corollary 2.1 and (9.1) yield

$$w^-(\mathbf{x}+\mathbf{y}, t) \geq w(\mathbf{x}+\mathbf{y}, t) - C\varepsilon = \lambda|\mathbf{y}|^2 + o(|\mathbf{y}|^2) - C\varepsilon \geq \frac{\lambda}{2}|\mathbf{y}|^2 - C\varepsilon = \varepsilon,$$

for $|\mathbf{y}|^2 = 2(C+1)\varepsilon/\lambda$ provided $|\mathbf{y}|$ is so small that satisfies $o(|\mathbf{y}|^2) \geq -\lambda|\mathbf{y}|^2/2$. Reasoning again as in Corollary 8.1 leads to $U_{\varepsilon,h,\tau}(\mathbf{x}+\mathbf{y}, t) = 1$ and thus to (9.5). $\square$

This result partially explains why it is feasible to approximate the time at which singularities occur, and in particular the extinction time. In view of (9.3), the approximate interface $\Sigma_{\varepsilon,h,\tau}(t)$ behaves much like the true one in that it either disappears locally in time, if $k = n$ (sphere), or at least it moves away from $\mathbf{x}$ when $2 \leq k < n$ (cylinder). Both (9.3) and (9.4) provide theoretical basis for the study of singularities via (4.2), which does not seem to be available for any other fully discrete method. Statement (9.5) quantifies a smearing effect of (4.2) observed in practice in the vicinity of singularities, which gives rise to a much more diffuse transition layer with thickness $\mathcal{O}(\varepsilon^{1/2})$ instead of $\mathcal{O}(\varepsilon)$.

To enhance resolution we may think of reducing locally the size of ε, along with that of h and τ, which would then become space-time dependent. This simple idea turns out to be extremely successful computationally [25], [28], [29], [32], but the ensuing fully discrete method is yet to be studied.

Acknowledgments. We would like to thank X.Y Chen, L.C. Evans, Y. Giga, and J. Spruck for enlightening discussions regarding properties of viscosity solutions.

Note added in proof. If y is caloric, namely $\partial_t y - \Delta y = 0$ in Q, $y = 0$ on $\partial\Omega \times (0,T)$, and $y^0 = y(\cdot, 0) = (-\Delta)^{-1} z^0$ with zero Dirichlet boundary condition and $z^0 \in L^\infty(\Omega)$, then

$$||y - \hat{H}_{h,\tau}(y)||_{L^\infty(Q)} \leq C(h^2 + \tau)|\log h|^2 ||z^0||_{L^\infty(\Omega)}$$

is valid in *any* space dimensions. This is a recent result proven in [37]. Whether this estimate is valid for noncaloric $y \in W_p^{2,1}(Q)$ is still an open problem.

REFERENCES

[1] S. Altschuler, S. Angenent, and Y. Giga, *Mean curvature flow through singularities for surfaces of rotation*, J. Geom. Anal., to appear.

[2] G. Barles, H.-M. Soner, and P.E. Souganidis, *Front propagation and phase field theory*, SIAM J. Control Optim., 31 (1993), pp. 439–469.

[3] G. Bellettini and M. Paolini, *Two examples of fattening for the curvature flow with a driving force*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. (9) Mat. Appl., 5 (1994), pp. 229–236.

[4] G. Bellettini, M. Paolini, and C. Verdi, *Γ-convergence of discrete approximations to interfaces with prescribed mean curvature*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. (9) Mat. Appl., 1 (1990), pp. 317–328.
[5] F. Brezzi and L.A. Caffarelli, *Convergence of the discrete free boundaries for finite element approximations*, RAIRO Modél. Math. Anal. Numér., 17 (1983), pp. 385–395.
[6] X. Chen and C.M. Elliott, *Asymptotics for a parabolic double obstacle problem*, Proc. Roy. Soc. London Ser. A, 444 (1994), pp. 429–445.
[7] Y.G. Chen, Y. Giga, and S. Goto, *Uniqueness and existence of viscosity solutions of generalized mean curvature flow equation*, J. Differential Geom., 33 (1991), pp. 749–786.
[8] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.
[9] P.G. Ciarlet and P.A. Raviart, *Maximum principle and uniform convergence for the finite element method*, Comput. Methods Appl. Mech. Engrg., 2 (1973), pp. 17–31.
[10] M.G. Crandall, H. Ishii, and P.-L. Lions, *User's guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc. (N.S.), 27 (1992), pp. 1–67.
[11] K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems II: optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$*, SIAM J. Numer. Anal., to appear.
[12] L.C. Evans, H.-M. Soner, and P.E. Souganidis, *Phase transitions and generalized motion by mean curvature*, Comm. Pure Appl. Math., 45 (1992), pp. 1097–1123.
[13] L.C. Evans and J. Spruck, *Motion of level sets by mean curvature. I*, J. Differential Geom., 33 (1991), pp. 635–681.
[14] A. Fetter, *L^∞-error estimate for an approximation of a parabolic variational inequality*, Numer. Math., 50 (1987), pp. 557–565.
[15] A. Friedman, *Variational Principles and Free Boundary Problems*, Wiley, New York, 1982.
[16] Y. Giga and S. Goto, *Geometric evolution of phase-boundaries*, in On the evolution of phase boundaries, IMA VMA 43 (M.E. Gurtin and G.B. MacFadden, eds.), Springer-Verlag, New York, 1992, pp. 51–65.
[17] Y. Giga, S. Goto, H. Ishii, and M.H. Sato, *Comparison principle and convexity preserving properties for singular degenerate parabolic equations on unbounded domains*, Indiana Univ. Math. J., 40 (1991), pp. 443–470.
[18] M. Grayson, *The heat equation shrinks embedded plane curves to round points*, J. Differential Geom., 26 (1987), pp. 285–314.
[19] G. Huisken, *Asymptotic behavior for singularities of the mean curvature flow*, J. Differential Geom., 31 (1990), pp. 285–299.
[20] T. Ilmanen, *Generalized flow of sets by mean curvature on a manifold*, Indiana Univ. Math. J., 41, 3 (1992), pp. 671–705.
[21] O.A. Ladyženskaja, V. Solonnikov, and N. Ural'ceva, *Linear and Quasilinear Equations of Parabolic Type*, Translation Mathematical Monographs 23, AMS, 1968.
[22] J.A. Nitsche, *L^∞-convergence of finite element approximations*, in Mathematical Aspects of Finite Element Methods LNM 606, Springer-Verlag, New York, 1977, pp. 261–274.
[23] R.H. Nochetto, *A note on the approximation of free boundaries by finite element methods*, RAIRO Modél. Math. Anal. Numér., 20 (1986), pp. 355–368.
[24] ———, *Error estimates for multidimensional singular parabolic problems*, Japan J. Appl. Math., 4 (1987), pp. 111–138.
[25] R.H. Nochetto, M. Paolini, S. Rovida, and C. Verdi, *Variational approximation of the geometric motion of fronts*, in Motion by Mean Curvature and Related Topics (G. Buttazzo and A. Visintin, eds.), Gruyter, Berlin, 1994, pp. 124–149.
[26] R.H. Nochetto, M. Paolini, and C. Verdi, *Optimal interface error estimates for the mean curvature flow*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), 21 (1994), pp. 193–212.
[27] ———, *Sharp error analysis for curvature dependent evolving fronts*, Math. Models Methods Appl. Sci., 3 (1993), pp. 771–723.
[28] ———, *Double obstacle formulation with variable relaxation parameter for smooth geometric front evolutions: asymptotic interface error estimates*, Asymptotic Anal., 10 (1995), pp. 173–198.
[29] ———, *A dynamic mesh method for curvature dependent evolving interfaces*, J. Comp. Phys., submitted.
[30] R.H. Nochetto and C. Verdi, *Convergence of double obstacle problems to the generalized geometric motion of fronts*, SIAM J. Math. Anal., to appear.
[31] S. Osher and J.A. Sethian, *Fronts propagating with curvature dependent speed: algorithms*

based on Hamilton-Jacobi formulations,, J. Comput. Phys., 79 (1988), pp. 12–49.

[32] M. PAOLINI AND C. VERDI, *Asymptotic and numerical analyses of the mean curvature flow with a space-dependent relaxation parameter*, Asymptotic Anal., 5 (1992), pp. 553–574.

[33] P. PIETRA AND C. VERDI, *On the convergence of the approximate free boundary for the parabolic obstacle problem*, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8), 79 (1985), pp. 159–171.

[34] M.H. PROTTER AND H.F. WEINBERGER, *Maximum Principles in Differential Equations*, Prentice Hall, Englewood Cliffs, 1967.

[35] R. RANNACHER AND R. SCOTT, *Some optimal error estimates for piecewise linear finite element approximations*, Math. Comp., 38 (1982), pp. 437–445.

[36] H.-M. SONER AND P.E. SOUGANIDIS, *Singularities and uniqueness of cylindrically symmetric surfaces moving by mean curvature*, Comm. Partial Differential Equations, 18 (1993), pp. 859–894.

[37] C. PALENCIA, *Maximum norm analysis of completely discrete finite element methods for parabolic problems*, SIAM J. Numer. Anal., to appear.